

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :  C12Q 1/68	A2	(11) International Publication Number: <b>WO 99/66075</b>  (43) International Publication Date: 23 December 1999 (23.12.99)
(21) International Application Number: PCT/US99/13426		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 14 June 1999 (14.06.99)		
(30) Priority Data: 09/098,079 15 June 1998 (15.06.98) US 09/097,889 15 June 1998 (15.06.98) US 09/302,681 30 April 1999 (30.04.99) US		
(71) Applicant ( <i>for all designated States except US</i> ): MITOKOR [US/US]; 11494 Sorrento Valley Road, San Diego, CA 92121 (US).		<b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
(72) Inventors; and		
(75) Inventors/Applicants ( <i>for US only</i> ): HERRNSTADT, Corinna [DE/US]; 3628 Ruette DeVille, San Diego, CA 92130 (US). GHOSH, Soumitra, S. [US/US]; 12334 Pathos Lane, San Diego, CA 92129 (US). CLEVINGER, William [US/US]; 1515 South Melrose Drive #43, Vista, CA 92083 (US). FAHY, Eoin, D. [US/US]; 7171 Canon Hill Way, San Diego, CA 92126 (US). DAVIS, Robert, E. [US/US]; 13272 Glencliff Way, San Diego, CA 92130 (US).		
(74) Agents: ROSENMAN, Stephen, J. et al.; Seed and Berry LLP, 6300 Columbia Center, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).		

(54) Title: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF EXTRAMITOCHONDRIAL DNA

(57) Abstract

Compositions and methods based on quantification of extramitochondrial DNA (exmtDNA) sequences are provided that are useful for detecting the presence of or risk for having a disease associated with altered mitochondrial function, and for identifying agents suitable for treating such diseases. The exmtDNA sequences have strong homology to authentic mitochondrial DNA (mtDNA) sequences.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF EXTRAMITOCHONDRIAL DNA

### TECHNICAL FIELD

The present invention relates generally to diseases in which altered mitochondrial function, such as free radical mediated oxidative injury, leads to tissue degeneration and, more specifically, to compositions and methods for detecting predisposition to such diseases by quantifying extramitochondrial DNA.

### BACKGROUND OF THE INVENTION

A number of degenerative diseases are thought to be caused by or be associated with alterations in mitochondrial function. These diseases include Alzheimer's Disease, diabetes mellitus, Parkinson's Disease, Huntington's disease, dystonia, Leber's hereditary optic neuropathy, schizophrenia, and myodegenerative disorders such as "mitochondrial encephalopathy, lactic acidosis, and stroke" (MELAS), and "myoclonic epilepsy ragged red fiber syndrome" (MERRF). Other diseases involving altered metabolism or respiration within cells may also be regarded as diseases associated with altered mitochondrial function.

Functional mitochondria contain gene products encoded by mitochondrial genes situated in mitochondrial DNA (mtDNA) and by extramitochondrial genes not situated in the circular mitochondrial genome. The 16.5 kb mtDNA encodes 22 tRNAs, two ribosomal RNAs (rRNA) and only 13 enzymes of the electron transport chain (ETC), the elaborate multi-complex mitochondrial assembly where, for example, respiratory oxidative phosphorylation takes place. The overwhelming majority of mitochondrial structural and functional proteins are encoded by extramitochondrial, and in most cases presumably nuclear, genes. Accordingly, mitochondrial and extramitochondrial genes may interact directly, or indirectly via gene products and their downstream intermediates, including metabolites, catabolites, substrates, precursors, cofactors and the like. Alterations in mitochondrial function, for example impaired electron transport activity, defective oxidative phosphorylation or

increased free radical production, may therefore arise as the result of defective mtDNA, defective extramitochondrial DNA, defective mitochondrial or extramitochondrial gene products, defective downstream intermediates or a combination of these and other factors.

5 Mitochondria are the subcellular organelles that manufacture bioenergetically essential adenosine triphosphate (ATP) by oxidative phosphorylation. Defective mitochondrial activity, including failure at any step of the ETC, may result in the generation of highly reactive free radicals that have the potential of damaging cells and tissues. These free radicals may include reactive oxygen species (ROS) such as  
10 superoxide, peroxynitrite and hydroxyl radicals, and potentially other reactive species that may be toxic to cells. For example, oxygen free radical induced lipid peroxidation is a well established pathogenetic mechanism in central nervous system (CNS) injury, such as that found in a number of degenerative diseases, and in ischemia (*i.e.*, stroke).

There are at least two deleterious consequences of exposure to reactive  
15 free radicals arising from mitochondrial dysfunction that adversely impact the mitochondria themselves. First, free radical mediated damage may inactivate one or more of the myriad proteins of the ETC. According to generally accepted theories of mitochondrial function, proper ETC respiratory activity requires maintenance of an electrochemical potential in the inner mitochondrial membrane by a coupled  
20 chemiosmotic mechanism. Free radical oxidative activity may dissipate this membrane potential, thereby preventing ATP biosynthesis and halting the production of a vital biochemical energy source. In addition, mitochondrial proteins such as cytochrome c and "apoptosis inducing factor" may leak out of the mitochondria after permeability transition and may induce the genetically programmed cell suicide sequence known as  
25 apoptosis or programmed cell death (PCD).

Second, free radical mediated damage may result in catastrophic mitochondrial collapse that has been termed "transition permeability". For example, rapid mitochondrial permeability transition likely entails changes in the inner mitochondrial transmembrane protein adenylate translocase that results in the formation  
30 of a "pore." In any event, because permeability transition is potentiated by free radical

exposure, it may be more likely to occur in the mitochondria of cells from patients having mitochondria associated diseases that are chronically exposed to such reactive free radicals.

Altered mitochondrial function characteristic of the mitochondria  
5 associated diseases may also be related to loss of mitochondrial membrane electrochemical potential by mechanisms other than free radical oxidation, and such transition permeability may result from direct or indirect effects of mitochondrial genes, gene products or related downstream mediator molecules and/or extramitochondrial genes, gene products or related downstream mediators, or from other known or  
10 unknown causes.

Diabetes mellitus is a common, degenerative disease affecting 5 to 10 percent of the population in developed countries. The propensity for developing diabetes mellitus is reportedly maternally inherited, suggesting a mitochondrial genetic involvement. (Alcolado, J.C. and Alcolado, R., *Br. Med. J.* 302:1178-1180 (1991);  
15 Reny, S.L., *International J. Epidem.* 23:886-890 (1994)). Diabetes is a heterogenous disorder with a strong genetic component; monozygotic twins are highly concordant and there is a high incidence of the disease among first degree relatives of affected individuals.

At the cellular level, the degenerative phenotype that may be  
20 characteristic of late onset diabetes mellitus includes indicators of altered mitochondrial respiratory function, for example impaired insulin secretion, decreased ATP synthesis and increased levels of reactive oxygen species. Studies have shown that diabetes mellitus may be preceded by or associated with certain related disorders. For example, it is estimated that forty million individuals in the U.S. suffer from late onset impaired  
25 glucose tolerance (IGT). IGT patients fail to respond to glucose with increased insulin secretion. A small percentage of IGT individuals (5-10%) progress to insulin deficient non-insulin dependent diabetes (NIDDM) each year. Some of these individuals further progress to insulin dependent diabetes mellitus (IDDM). These forms of diabetes mellitus, NIDDM and IDDM, are associated with decreased release of insulin by  
30 pancreatic beta cells and/or a decreased end-organ response to insulin. Other symptoms

of diabetes mellitus and conditions that precede or are associated with diabetes mellitus include obesity, vascular pathologies, peripheral and sensory neuropathies, blindness and deafness.

Parkinson's disease (PD) is a progressive, neurodegenerative disorder  
5 associated with altered mitochondrial function and characterized by the loss and/or atrophy of dopamine-containing neurons in the *pars compacta* of the *substantia nigra* of the brain. Like Alzheimer's Disease (AD), PD also afflicts the elderly. It is characterized by bradykinesia (slow movement), rigidity and a resting tremor. Although L-Dopa treatment reduces tremors in most patients for a while, ultimately the  
10 tremors become more and more uncontrollable, making it difficult or impossible for patients to even feed themselves or meet their own basic hygiene needs.

It has been shown that the neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) induces parkinsonism in animals and man at least in part through its effects on mitochondria. MPTP is converted to its active metabolite, MPP+,  
15 in dopamine neurons; it then becomes concentrated in the mitochondria. The MPP+ then selectively inhibits the mitochondrial enzyme NADH:ubiquinone oxidoreductase ("Complex I"), leading to the increased production of free radicals, reduced production of adenosine triphosphate, and ultimately, the death of affected dopamine neurons.

Mitochondrial Complex I is composed of 40-50 subunits; most are  
20 encoded by the nuclear genome and seven by the mitochondrial genome. Since parkinsonism may be induced by exposure to mitochondrial toxins that affect Complex I activity, it appears likely that defects in Complex I proteins may contribute to the pathogenesis of PD by causing a similar biochemical deficiency in Complex I activity. Indeed, defects in mitochondrial Complex I activity have been reported in the  
25 blood and brain of PD patients (Parker et al., *Am. J. Neurol.* 26:719-723, 1989).

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that is characterized by loss and/or atrophy of neurons in discrete regions of the brain, and that is accompanied by extracellular deposits of  $\beta$ -amyloid and the intracellular accumulation of neurofibrillary tangles. It is a uniquely human disease, affecting over  
30 13 million people worldwide. It is also a uniquely tragic disease. Many individuals

who have lived normal, productive lives are slowly stricken with AD as they grow older, and the disease gradually robs them of their memory and other mental faculties. Eventually, they cease to recognize family and loved ones, and they often require continuous care until their eventual death.

- 5        There is evidence that defects in oxidative phosphorylation within the mitochondria are at least a partial cause of sporadic AD. The enzyme cytochrome c oxidase (COX), which makes up part of the mitochondrial electron transport chain (ETC), is present in normal amounts in AD patients; however, the catalytic activity of this enzyme in AD patients and in the brains of AD patients at autopsy has been found  
10      to be abnormally low. This suggests that the COX in AD patients is defective, leading to decreased catalytic activity that in some fashion causes or contributes to the symptoms that are characteristic of AD.

Focal defects in energy metabolism in the mitochondria, with accompanying increases in oxidative stress, may be associated with AD. It is well-established that energy metabolism is impaired in AD brain (Palmer et al., *Brain Res.* 645:338-42, 1994; Pappolla et al., *Am. J. Pathol.* 140:621-28, 1992; Jeandel et al., *Gerontol.* 35:275, 1989; Balazs et al., *Neurochem. Res.* 19:1131-37, 1994; Mecocci et al., *Ann. Neurol.* 36:747-751, 1994; Gsell et al., *J. Neurochem.* 64:1216-23, 1995). For example, regionally specific deficits in energy metabolism in AD brains have been  
15      reported in a number of positron emission tomography studies (Kuhl, et al., *J. Cereb. Blood Flow Metab.* 7:S406, 1987; Grady, et al., *J. Clin. Exp. Neuropsychol.* 10:576-96, 1988; Haxby et al., *Arch. Neurol.* 47:753-60, 1990; Azari et al., *J. Cereb. Blood Flow Metab.* 13:438-47, 1993). Metabolic defects in the temporoparietal neocortex of AD patients apparently presage cognitive decline by several years. Skin fibroblasts from  
20      AD patients display decreased glucose utilization and increased oxidation of glucose, leading to the formation of glycosylation end products (Yan et al., *Proc. Nat. Acad. Sci. USA* 91:7787-91, 1994). Cortical tissue from postmortem AD brain shows decreased activity of the mitochondrial enzymes pyruvate dehydrogenase (Sheu et al., *Ann. Neurol.* 17:444-49, 1985) and  $\alpha$ -ketoglutarate dehydrogenase (Mastrogiacomo et al., *J. Neurochem.* 6:2007-14, 1994), which are both key enzymes in energy metabolism.  
25  
30

Functional magnetic resonance spectroscopy studies have shown increased levels of inorganic phosphate relative to phosphocreatine in AD brain, suggesting an accumulation of precursors that arises from decreased ATP production by mitochondria (Pettegrew et al., *Neurobiol. of Aging* 15:117-32, 1994; Pettigrew et al., *Neurobiol. of Aging* 16:973-75, 1995). In addition, the levels of pyruvate, but not of glucose or lactate, are reported to be increased in the cerebrospinal fluid of AD patients, consistent with defects in cerebral mitochondrial electron transport chain (ETC) activity (Parnetti et al., *Neurosci. Lett.* 199:231-33, 1995).

Signs of oxidative injury are prominent features of AD pathology and, as noted above, reactive oxygen species (ROS) are critical mediators of neuronal degeneration. Indeed, studies at autopsy show that markers of protein, DNA and lipid peroxidation are increased in AD brain (Palmer et al., *Brain Res.* 645:338-42, 1994; Pappolla et al., *Am. J. Pathol.* 140:621-28, 1992; Jeandel et al., *Gerontol.* 35:275-82, 1989; Balazs et al., *Arch. Neurol.* 4:864, 1994; Mecocci et al., *Ann. Neurol.* 36:747-51, 1994; Smith et al., *Proc. Nat. Acad. Sci. USA* 88:10540-43, 1991). In hippocampal tissue from AD but not from controls, carbonyl formation indicative of protein oxidation is increased in neuronal cytoplasm, and nuclei of neurons and glia (Smith et al., *Nature* 382:120-21, 1996). Neurofibrillary tangles also appear to be prominent sites of protein oxidation (Schweers et al., *Proc. Nat. Acad. Sci. USA* 92:8463, 1995; Blass et al., *Arch. Neurol.* 4:864, 1990). Under stressed and non-stressed conditions incubation of cortical tissue from AD brains taken at autopsy demonstrate increased free radical production relative to non-AD controls. In addition, the activities of critical antioxidant enzymes, particularly catalase, are reduced in AD (Gsell et al., *J. Neurochem.* 64:1216-23, 1995), suggesting that the AD brain is vulnerable to increased ROS production. Thus, oxidative stress may contribute significantly to the pathology of mitochondria associated diseases such as AD, where mitochondrial dysfunction and/or elevated ROS may be present.

One hallmark pathology of AD is the death of selected neuronal populations in discrete regions of the brain. Cell death in AD is presumed to be apoptotic because signs of programmed cell death (PCD) are seen and indicators of

active gliosis and necrosis are not found. (Smale et al., *Exp. Neurolog.* 133:225-230, 1995; Cotman et al., *Molec. Neurobiol.* 10:19-45, 1995.) The consequences of cell death in AD, neuronal and synaptic loss, are closely associated with the clinical diagnosis of AD and are highly correlated with the degree of dementia in AD (DeKosky et al., *Ann. Neurology* 27:457-464, 1990).

Mitochondrial dysfunction is thought to be critical in the cascade of events leading to apoptosis in various cell types (Kroemer et al., *FASEB J.* 9:1277-87, 1995), and may be a cause of apoptotic cell death in neurons of the AD brain. Altered mitochondrial physiology may be among the earliest events in PCD (Zamzami et al., *J. Exp. Med.* 182:367-77, 1995; Zamzami et al., *J. Exp. Med.* 181:1661-72, 1995) and elevated reactive oxygen species (ROS) levels that result from such altered mitochondrial function may initiate the apoptotic cascade (Ausserer et al., *Mol. Cell. Biol.* 14:5032-42, 1994). In several cell types, including neurons, reduction in the mitochondrial membrane potential ( $\Delta\Psi_m$ ) precedes the nuclear DNA degradation that accompanies apoptosis. In cell-free systems, mitochondrial, but not nuclear, enriched fractions are capable of inducing nuclear apoptosis (Newmeyer et al., *Cell* 70:353-64, 1994). Perturbation of mitochondrial respiratory activity leading to altered cellular metabolic states, such as elevated intracellular ROS, may occur in mitochondria associated diseases and may further induce pathogenetic events via apoptotic mechanisms.

Oxidatively stressed mitochondria may release a pre-formed soluble factor that can induce chromosomal condensation, an event preceding apoptosis (Marchetti et al., *Cancer Res.* 56:2033-38, 1996). In addition, members of the Bcl-2 family of anti-apoptosis gene products are located within the outer mitochondrial membrane (Monaghan et al., *J. Histochem. Cytochem.* 40:1819-25, 1992) and these proteins appear to protect membranes from oxidative stress (Korsmeyer et al., *Biochim. Biophys. Act.* 1271:63, 1995). Localization of Bcl-2 to this membrane appears to be indispensable for modulation of apoptosis (Nguyen et al., *J. Biol. Chem.* 269:16521-24, 1994). Thus, changes in mitochondrial physiology may be important mediators of apoptosis. To the extent that apoptotic cell death is a prominent feature of neuronal loss

in AD, mitochondrial dysfunction may be critical to the progression of this disease and may also be a contributing factor in other mitochondria associated diseases.

Regardless of whether a defect underlying a disease associated with altered mitochondrial function may have mitochondrial or extramitochondrial origins, 5 and regardless of whether a defect underlying altered mitochondrial function has been identified, the present invention provides methods that are useful for determining the risk or presence of diseases associated with such altered mitochondrial function, and for identifying agents that are suitable for treating such diseases. In particular, as is elaborated herein below, the present invention provides compositions and methods for 10 the detection of diseases associated with altered mitochondrial function by quantification of unusual mtDNA-like sequences not found in mitochondria and referred to as extramitochondrial DNA (exmtDNA), and other related advantages.

#### SUMMARY OF THE INVENTION

Briefly stated, the present invention is directed to compositions and 15 methods useful for detecting mitochondria associated diseases and involving extramitochondrial DNA (exmtDNA) sequences that are highly homologous to mitochondrial DNA (mtDNA). In one aspect the invention provides a method for determining the risk for or presence of a disease associated with altered mitochondrial function in a first subject suspected of having or being at risk for having such a disease, 20 by comparing a ratio  $r$  for each of a first and a second biological sample containing extramitochondrial DNA and mitochondrial DNA, the first biological sample being obtained from the first subject and the second sample being obtained from a second subject known to be free of a risk or presence of a disease associated with altered mitochondrial function, using the formula:

25

$$r = x/(x + y)$$

wherein  $x$  is the amount of extramitochondrial DNA in a sample, and  $y$  is the amount of mitochondrial DNA in the sample; and therefrom determining the risk or presence of 30 the disease. In an embodiment of the invention, the ratio  $r$  is calculated by a method

that comprises contacting a biological sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in the extramitochondrial DNA and present in the mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of  
5 the primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, in order to therefrom quantify the extramitochondrial DNA and the mitochondrial DNA.

In another embodiment, the ratio  $r$  is calculated by a method comprising  
10 contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in the amplified extramitochondrial DNA and present in the amplified mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA and to the  
15 mitochondrial DNA; and detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA. In another embodiment of this aspect of the invention the ratio  $r$  is calculated by a method comprising contacting a biological sample containing extramitochondrial DNA and mitochondrial DNA with an  
20 oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in the extramitochondrial DNA and present in the mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a  
25 first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from the first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA.

In another embodiment of this aspect of the invention the ratio  $r$  is calculated by a method comprising contacting a sample containing amplified  
30 extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer

having a nucleotide sequence that is complementary to a sequence present in the amplified extramitochondrial DNA and present in the amplified mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization  
5 and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from the first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA.

In another embodiment of this aspect of the invention the biological  
10 sample is treated by heating it in water to lyse cells contained in the sample, and then extracting cellular DNA from the lysed cells using an aqueous DNA extraction procedure. In certain embodiments of the invention the sample comprises a crude buffy coat fraction of whole blood. In certain other embodiments of the invention, the method further comprises the step of determining the ApoE genotype of the first subject and  
15 correlating said genotype with the risk or presence of disease. In some embodiments of the invention, the disease associated with altered mitochondrial function may be Alzheimer's Disease, Huntington's Disease, Parkinson's Disease, dystonia, schizophrenia, non-insulin dependent diabetes mellitus, mitochondrial encephalopathy, lactic acidosis, and stroke, myoclonic epilepsy ragged red fiber syndrome, and Leber's  
20 hereditary optic neuropathy.

Another aspect of the invention provides a method for quantifying extramitochondrial DNA, comprising: contacting a sample containing extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence complementary to at least a portion of the extramitochondrial DNA under conditions  
25 and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA; and detecting hybridization of the primer to the extramitochondrial DNA, and therefrom quantifying the extramitochondrial DNA.

It is another aspect of the invention to provide a method for quantifying extramitochondrial DNA, comprising: contacting a sample containing  
30 extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence

complementary to at least a portion of the extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA; and detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a product, and therefrom quantifying the extramitochondrial DNA.

- 5 Another aspect of the invention provides a method for quantifying extramitochondrial DNA, comprising: contacting a sample containing amplified extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence complementary to at least a portion of the extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial  
10 DNA; and detecting hybridization of the primer to the extramitochondrial DNA, therefrom quantifying the extramitochondrial DNA.

- In yet another aspect of the invention, a method is provided for quantifying extramitochondrial DNA by contacting a sample containing amplified extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence  
15 complementary to at least a portion of the extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA; and detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a product, and therefrom quantifying the extramitochondrial DNA.

- In one embodiment the extramitochondrial DNA is amplified by  
20 polymerase chain reaction, transcriptional amplification systems or self-sustained sequence replication. In certain embodiments of the various aspects of the invention, a single oligonucleotide primer is used. In certain embodiments of the invention a primer extension assay is used. In certain embodiments of the invention, the step of detecting may be by polymerase chain reaction, primer extension assay, ligase chain reaction or  
25 restriction fragment length polymorphism analysis.

- In certain embodiments of the invention, the amount of extramitochondrial DNA in a biological sample is quantified by determining the presence in the sample of a nucleotide sequence that may be SEQ ID NO:1, a portion of SEQ ID NO:1, SEQ ID NO:3, a portion of SEQ ID NO:3, an extramitochondrial DNA  
30 sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of

SEQ ID NO:2 and (ii) contains at least one nucleotide substitution of Figure 4 at a corresponding nucleotide position, or an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one nucleotide deletion of Figure 4 at a corresponding nucleotide 5 position. In one embodiment the portion of the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial cytochrome c oxidase. In another embodiment the portion of SEQ ID NO:1 corresponds to a portion of a mitochondrial cytochrome c oxidase encoding sequence that may be portion of a cytochrome c oxidase 1 (CO1) encoding sequence or 10 a portion of a cytochrome c oxidase 2 (CO2) encoding sequence. In still other embodiments, the portion of the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial ATP synthetase subunit. In other embodiments, the portion of SEQ ID NO:1 corresponds to a portion of a mitochondrial ATP synthetase subunit encoding sequence that may be a 15 portion of a sequence encoding ATP synthetase subunit 6 or a portion of a sequence encoding ATP synthetase subunit 8.

In some embodiments the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of SEQ ID NO:2 that may be a portion of a sequence encoding ND1, a sequence encoding a portion of ND2 or a sequence encoding a portion of CO3. 20 In other embodiments, the portion of the nucleotide sequence of SEQ ID NO:3 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial ATP synthetase subunit, which in some embodiments may further be a portion of a sequence encoding ATP synthetase subunit 6 or a portion of a sequence encoding ATP synthetase subunit 8. In still other embodiments, the nucleotide 25 sequence of SEQ ID NO:1 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial tRNA, while in yet other embodiments the portion of the nucleotide sequence of SEQ ID NO:3 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial tRNA.

In another aspect the invention provides an isolated nucleic acid 30 comprising all or a portion of the nucleotide sequence of SEQ ID NO:1 or a

complementary sequence thereto. In another aspect the invention provides an isolated nucleic acid comprising all or a portion of a nucleotide sequence of SEQ ID NO:1 or a complementary sequence thereto, wherein the sequence of the isolated nucleic acid differs by at least one nucleotide from the corresponding sequence of a nucleic acid

5 comprising the nucleotide sequence of SEQ ID NO:2 or a complementary sequence thereto. In another aspect the invention provides an isolated nucleic acid comprising all or a portion of the nucleotide sequence of SEQ ID NO:3 or a complementary sequence thereto. In another aspect the invention provides an isolated nucleic acid comprising all or a portion of a nucleotide sequence of SEQ ID NO:3 or a complementary sequence

10 thereto, wherein the sequence of the isolated nucleic acid differs by at least one nucleotide from the corresponding sequence of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:2 or a complementary sequence thereto.

In another aspect the invention provides a method for determining the risk or presence of a disease associated with altered mitochondrial function in a subject

15 suspected of having or being at risk for having such a disease, by quantifying the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from the subject, and therefrom determining the risk or presence of the disease. It is another aspect of the invention to provide a method for determining the risk or presence of a disease associated with altered mitochondrial function in a first

20 subject suspected of having or being at risk for having such a disease, by comparing the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from the first subject to the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from a second subject, and therefrom determining the risk or presence of the disease. In another aspect the

25 invention provides a method for determining the risk or presence of a disease associated with altered mitochondrial function in a first subject suspected of having or being at risk for having such a disease, by quantifying the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from the subject and comparing the amount of extramitochondrial DNA and the amount of mitochondrial

30 DNA to the amount of extramitochondrial DNA and the amount of mitochondrial DNA

in a biological sample from a second subject known to be free of a risk or presence of a disease associated with altered mitochondrial function, and therefrom determining the risk or presence of the disease.

Another aspect of the invention provides a method of monitoring a telomere by detecting in a telomeric region the presence of a nucleic acid molecule comprising all or a portion of SEQ ID NO:1 or a complementary portion thereto. In one embodiment, the detected nucleic acid molecule comprises an exmtDNA sequence. In another aspect, the invention provides a method of monitoring a telomere by detecting in a telomeric region the presence of a nucleic acid molecule comprising all or a portion of SEQ ID NO:3 or a complementary portion thereto.

Turning to another aspect, the invention provides a method of identifying an agent suitable for treating a disease associated with altered mitochondrial function, by comparing a ratio  $r$  from a sample obtained before contacting a biological source with a candidate agent to the ratio  $r$  from a sample obtained after contacting the biological source with the candidate agent, said ratio  $r$  calculated using the formula:

$$r = x/(x + y)$$

wherein  $x$  is the amount of extramitochondrial DNA in a sample, and  $y$  is the amount of mitochondrial DNA in the sample; and therefrom determining the suitability of said candidate agent for treating a disease associated with altered mitochondrial function. In one embodiment, the biological sample may be a crude buffy coat fraction of whole blood. In another embodiment, the biological sample is treated by heating in water to lyse cells contained in the sample, and then extracting cellular DNA from lysed cells using an aqueous DNA extraction procedure. In another embodiment, the ratio  $r$  is calculated by contacting a sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in the extramitochondrial DNA and present in the mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial

DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio  $r$ .

In another embodiment of the invention, the ratio  $r$  is calculated by contacting a sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in the extramitochondrial DNA and present in the mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from the first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio. In another embodiment, the ratio  $r$  is calculated by contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said amplified extramitochondrial DNA and present in said amplified mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio  $r$ .

In yet another embodiment, the ratio  $r$  is calculated by contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in the amplified extramitochondrial DNA and present in the amplified mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from said first product,

and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio.

In another embodiment of the invention, comparing the ratio  $r$  from a sample obtained before contacting a biological source with a candidate agent to the ratio 5  $r$  from a sample obtained after contacting the biological source with the candidate agent comprises determination of the presence in the sample of a nucleotide sequence of SEQ ID NO:1 or portion thereof, or a nucleotide sequence of SEQ ID NO:3 or a portion thereof, or an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one 10 nucleotide substitution of Figure 4 at a corresponding nucleotide position, or an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one nucleotide deletion of Figure 4 at a corresponding nucleotide position. In another embodiment, the nucleotide sequence of SEQ ID NO:1 or a portion thereof corresponds 15 to a mitochondrial cytochrome c oxidase encoding sequence of SEQ ID NO:2 or a portion thereof. In another embodiment, the mitochondrial cytochrome c oxidase encoding sequence of SEQ ID NO:2 or a portion thereof is a sequence encoding CO I or a portion thereof, or a sequence encoding CO2 or a portion thereof. In another embodiment, the nucleotide sequence of SEQ ID NO:1 or portion thereof, or the 20 nucleotide sequence of SEQ ID NO:3 or portion thereof corresponds to a mitochondrial ATP synthetase subunit encoding sequence of SEQ ID NO:2 or a portion thereof. In another embodiment, the mitochondrial ATP synthetase subunit encoding sequence of SEQ ID NO:2 or a portion thereof may be a sequence encoding ATP synthetase subunit 6 or a portion thereof, or a sequence encoding ATP synthetase subunit 8 or a portion 25 thereof. In another embodiment, the nucleotide sequence of SEQ ID NO:1 corresponds to a sequence of SEQ ID NO:2 or a portion thereof that may be a sequence encoding a truncated NADH dehydrogenase subunit 1 or a portion thereof, a sequence encoding NADH dehydrogenase subunit 2 or a portion thereof or a sequence encoding truncated CO3 or a portion thereof.

In other embodiments of the invention, the disease associated with altered mitochondrial function may be Alzheimer's Disease, Huntington's Disease, Parkinson's Disease, dystonia, schizophrenia, non-insulin dependent diabetes mellitus, mitochondrial encephalopathy, lactic acidosis, and stroke, myoclonic epilepsy ragged red fiber syndrome, or Leber's hereditary optic neuropathy.

In another aspect, the invention provides a method of identifying an agent suitable for treating a subject suspected of being at risk for having a disease associated with altered mitochondrial function, by determining the apolipoprotein E genotype of the subject; comparing a ratio  $r$  in a biological sample obtained from the subject before contacting the sample with a candidate agent to the ratio  $r$  in a biological sample obtained from the subject after contacting the sample with a candidate agent, the ratio  $r$  calculated using the formula:

$$r = x / (x + y)$$

15

wherein  $x$  is the amount of extramitochondrial DNA in the sample, and  $y$  is the amount of mitochondrial DNA in the sample; and therefrom determining the suitability of said candidate agent for treating the disease associated with altered mitochondrial function. In another embodiment, the disease associated with altered mitochondrial function is

20 Alzheimer's disease.

It is another aspect of the invention to provide a method of correlating a ratio  $r$  with the suitability of an agent for treating Alzheimer's disease in a subject, by determining a ratio  $r$  in a biological sample obtained from the subject, said ratio  $r$  calculated using the formula:

25

$$r = x / (x + y)$$

wherein  $x$  is the amount of extramitochondrial DNA in the sample, and  $y$  is the amount of mitochondrial DNA in the sample; contacting said subject with a candidate agent and

30 evaluating the subject for alterations in the AD disease state, and therefrom correlating the suitability of the agent for treating AD in the subject with  $r$ . In another embodiment, the apolipoprotein E genotype of the subject is determined, and therefrom

the suitability of the agent for treating AD in the subject is correlated with  $r$  and with the apolipoprotein E genotype.

These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings. In addition, 5 various references are set forth herein which describe in more detail certain aspects of this invention, and are therefore incorporated by reference in their entirety.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the nucleotide sequence of SEQ ID NO:1.

Figure 2 depicts the nucleotide sequence of SEQ ID NO:2.

10 Figure 3 depicts the nucleotide sequence of SEQ ID NO:3.

Figure 4 depicts human extramitochondrial DNA nucleotide substitutions and deletions relative to the corresponding human mtDNA sequence of SEQ ID NO:2.

15 Figure 5 shows radiation hybrid chromosomal mapping of an extramitochondrial DNA sequence.

Figure 6 depicts detection of mtDNA and exmtDNA by Southern blot hybridization.

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed generally to compositions and methods 20 for diagnosing the risk or presence of a disease associated with altered mitochondrial function in a subject, and to compositions and methods for the identification of agents that may be suitable for treating a disease associated with altered mitochondrial function. The invention utilizes a ratio,  $r$ , that may be useful for pharmacogenomic purposes, for example to stratify patient populations according to the suitability of 25 particular therapeutic agents for use in such populations. The ratio  $r$  is the ratio of the amount of exmtDNA in a biological sample relative to the sum of the amount of exmtDNA plus mtDNA in the sample. As expressed quantitatively, the ratio  $r$  may be calculated using the formula:

$$r = x/(x + y)$$

wherein

x is the amount of exmtDNA in a sample, and

5 y is the amount of mtDNA in the sample.

In various aspects of the invention, as elaborated more fully herein, quantification of x and y provide, through calculation of r, parameters useful in diagnosis of a disease associated with altered mitochondrial function and in screening assays for agents that may be suitable for the treatment of such a disease.

10 As discussed above, "altered mitochondrial function" may refer to any condition or state, including those that accompany a disease, where any structure or activity that is directly or indirectly related to a mitochondrial function has been changed. Altered mitochondrial function may have its origin in extramitochondrial structures or events as well as in mitochondrial structures or events, in direct interactions between mitochondrial and extramitochondrial genes and/or their gene products, or in structural or functional changes that occur as the result of interactions between intermediates that may be formed as the result of such interactions, including metabolites, catabolites, substrates, precursors, cofactors and the like.

Also as discussed above, altered mitochondrial function may include  
20 (but need not be limited to) altered respiratory or metabolic activity in some or all cells of a biological source. For example, markedly impaired ETC activity may be an example of altered mitochondrial function, as may be generation of increased ROS or defective oxidative phosphorylation. As further examples, altered mitochondrial membrane potential, induction of apoptotic pathways and formation of atypical  
25 chemical and biochemical crosslinked species within a cell, whether by enzymatic or non-enzymatic mechanisms, may all be regarded as indicative of altered mitochondrial function. Without wishing to be bound by theory, alterations in the ratio r are believed to reflect chemical changes within affected cells that quantitatively influence recoveries of exmtDNA and/or mtDNA. For example, biochemical crosslinking events may result  
30 in the formation of DNA aggregates, DNA adducts or other molecular species that affect DNA recoveries following extraction procedures.

According to the present invention, alterations in the ratio  $r$  as defined above provide a novel and useful parameter for diagnosing the risk or presence of a disease associated with altered mitochondrial function in a subject, and for identifying agents that may be suitable for treating a disease associated with altered mitochondrial function. As discussed above, a number of diseases, including several degenerative diseases, are associated with alterations in mitochondrial function. Further, detection of an appropriate parameter of altered mitochondrial function can provide preclinical evidence for a risk of or predisposition to a disease.

Determination of the ratio  $r$  involves quantification of exmtDNA ( $x$ ) and mtDNA ( $y$ ) that may be based on strong but not necessarily absolute nucleotide sequence conservation when corresponding portions of mtDNA and exmtDNA are compared, as discussed herein. In most preferred embodiments of the invention, determination of  $r$  is accomplished by detecting minor nucleotide sequence differences in highly conserved mtDNA and exmtDNA regions, as elaborated below. The invention provides compositions and methods that include the use of nucleic acid molecules, or portions thereof, having nucleotide sequences that are found in the human mtDNA sequence SEQ ID NO:2 (Anderson et al., *Nature* 290:457, 1981) and fragments of SEQ ID NO:2 that are suitable for use as oligonucleotide primers in nucleic acid primer extension or amplification techniques, as hybridization probes for the detection of complementary nucleotide sequences in a sample or for any number of additional uses that are well known to those familiar with the art. ExmtDNA may be nuclear DNA, including chromosomal and non-chromosomal DNA, or non-nuclear extramitochondrial DNA that may be from any subcellular compartment, provided it is not mtDNA.

Nucleic acid sequences within the scope of the invention include isolated DNA and RNA sequences that specifically hybridize under conditions of moderate or high stringency to exmtDNA nucleotide sequences, including exmtDNA sequences disclosed herein or fragments thereof, and their complements. As used herein, conditions of moderate stringency, as known to those having ordinary skill in the art, and as defined by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed.

Vol. 1, pp. 1.101-104. Cold Spring Harbor Laboratory Press (1989), include use of a prewashing solution for the nitrocellulose filters 5X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0), hybridization conditions of 50% formamide, 6X SSC at 42°C (or other similar hybridization solution), and washing conditions of about 50-60°C, 0.5X SSC, 0.1% SDS. Conditions of high stringency are defined as hybridization conditions as above, and with washing at 60-68°C, 0.2X SSC, 0.1% SDS. In other embodiments, hybridization to an exmtDNA nucleotide sequence may be at normal stringency, which is approximately 25-30°C below Tm of the native duplex (e.g., 5X SSPE, 0.5% SDS, 5X Denhardt's solution, 50% formamide, at 42°C or equivalent conditions), at low stringency hybridizations, which utilize conditions approximately 40°C below Tm, or at high stringency hybridizations, which utilize conditions approximately 10°C below Tm. The skilled artisan will recognize that the temperature, salt concentration, and chaotrope composition of hybridization and wash solutions may be adjusted as necessary according to factors such as the length and nucleotide base composition of the probe.

15 (See also, e.g., Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing, 1987.)

An "isolated nucleic acid molecule" refers to a polynucleotide molecule in the form of a separate fragment or as a component of a larger nucleic acid construct, that has been separated from its source cell (including the chromosome it normally resides in) at least once, preferably in a substantially pure form. Isolated nucleic acids may be nucleic acids having particular disclosed nucleotide sequences or may be regions, portions or fragments thereof. Those having ordinary skill in the art are able to prepare isolated nucleic acids having the complete nucleotide sequence, or the sequence of any portion of a particular isolated nucleic acid molecule, when provided with the appropriate nucleic acid sequence information as disclosed herein. Nucleic acid molecules may be comprised of a wide variety of nucleotides, including DNA, RNA, nucleotide analogues such as phosphorothioates or peptide nucleic acids, or other analogues with which those skilled in the art will be familiar, or some combination of these.

The present invention, as described herein, provides exmtDNA sequences and isolated exmtDNA nucleic acid molecules. exmtDNA may be isolated from genomic DNA, typically by first generating an appropriate DNA library through techniques for constructing libraries that are known in the art (see Sambrook et al., 5 *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, 1989) or purchased from commercial sources (e.g., Clontech, Palo Alto, California). Briefly, genomic DNA libraries can be constructed in chromosomal vectors, such as YACs (yeast artificial chromosomes), bacteriophage vectors, such as pBeloBAC11, λEMBL3, λgt10, cosmids, or plasmids. Alternatively, isolated exmtDNA may be prepared by 10 preferentially amplifying exmtDNA sequences present in biological samples using, for example, DNA amplification methodologies such as PCR or other amplification techniques that are well known in the art, with suitable oligonucleotide primers complementary to exmtDNA sequences as disclosed herein.

In one embodiment, known mtDNA sequences derived from SEQ ID 15 NO:2 (Anderson et al., *Nature* 290:457, 1981) may be utilized to design oligonucleotide hybridization probes suitable for screening genomic libraries. Preferably, such oligonucleotide probes are 18-30 bases in length and have sequences that, under the hybridization conditions selected, hybridize to complementary exmtDNA sequences lacking nucleotide substitutions, insertions or deletions ("mutations") relative to the 20 corresponding region of the mtDNA sequence of SEQ ID NO:2.

Portions of an exmtDNA sequence and the mtDNA sequence of SEQ ID NO:2 are regarded as "corresponding" nucleic acid sequences, regions, fragments or the like, based on the convention for numbering mtDNA nucleic acid positions according to SEQ ID NO:2 (Anderson et al., *Nature* 290:457, 1981), wherein an exmtDNA sequence 25 is aligned with the mtDNA sequence of SEQ ID NO:2 such that at least 70%, preferably at least 80% and more preferably at least 90% of the nucleotides in a given sequence of at least 20 consecutive nucleotides of a sequence are identical. In certain preferred embodiments, an exmtDNA sequence is greater than 95% identical to a corresponding mtDNA sequence. In certain particularly preferred embodiments, an exmtDNA 30 sequence is identical to a corresponding mtDNA sequence. Those oligonucleotide

probes having sequences that are identical in corresponding regions of mtDNA and exmtDNA may be identified and selected following hybridization target DNA sequence analysis, to verify the absence of mutations in the target exmtDNA sequence relative to the primer mtDNA-derived sequence.

- 5 To facilitate hybridization detection, the oligonucleotide may be conveniently labeled, generally at the 5' end, with a reporter molecule, such as a radionuclide, e.g.,  $^{32}\text{P}$ , enzymatic label, protein label, fluorescent label, biotin or other suitable labeling moieties known in the art. Such libraries are then generally plated as phage or colonies, depending upon the vector used. Subsequently, a plate replica to  
10 which the colonies or phage have been transferred, such as a nitrocellulose or nylon membrane or the like, is probed to identify candidate clones that contain the exmtDNA sequence. Such candidates may be verified as containing exmtDNA by any of various means including, for example, DNA sequence analysis or hybridization with a second, non-overlapping probe selected as described above to hybridize with target exmtDNA  
15 sequences lacking nucleotide substitutions, deletions or insertions relative to the corresponding portion of the mtDNA sequence of SEQ ID NO:2.

- Once a library is identified as containing exmtDNA, the exmtDNA can be isolated by amplification. Briefly, when using genomic library DNA as a template, amplification primers are designed based upon known mtDNA sequences (SEQ ID  
20 NO:2) and primer "walking" is used to select primers that anneal to exmtDNA regions that are identical to mtDNA sequences. The primers preferably have a GC content of about 50% and contain restriction sites to facilitate cloning. Primers do not have self-complementary sequences, nor do they contain complementary sequences at their 3' end (to prevent primer-dimer formation). The primers are annealed to genomic DNA and  
25 sufficient amplification cycles are performed to yield a product readily visualized by gel electrophoresis and staining. The amplified fragment is purified and inserted into a vector, such as  $\lambda\text{gt}10$  or pBS(M13+), and propagated. Confirmation of the nature of the fragment is obtained by DNA sequence analysis.

- As an example of detection of mtDNA-like sequences in a DNA library,  
30 an oligonucleotide having a nucleotide sequence present in a portion of any human

mtDNA gene, preferably one of the human mtDNA encoded genes NADH dehydrogenase subunit 1 (ND1), NADH dehydrogenase subunit 2 (ND2), NADH dehydrogenase subunit 3 (ND3), NADH dehydrogenase subunit 4 (ND4), NADH dehydrogenase subunit 4L (ND4L), NADH dehydrogenase subunit 5 (ND5), NADH dehydrogenase subunit 6 (ND6), tRNA<sup>Lys</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Gln</sup>, tRNA<sup>Met</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>Asp</sup> or cytochrome c oxidase 3 (CO3) and more preferably one of the human mtDNA encoded genes CO1, CO2, ATPase 8 or ATPase 6, may be labeled and used as a probe on a human genomic DNA library. An initial hybridization at normal stringency may yield candidate clones or fragments. If no hybridization is initially observed, varying degrees of stringency may be used. (See Sambrook et al., Ausubel et al., *supra*, and other well-known sources for stringency conditions.)

Where it is advantageous to use oligonucleotide primers according to the present invention, such primers may be 10-60 nucleotides in length, preferably 15-35 nucleotides and still more preferably 18-30 nucleotides in length. Primers as described above for use in isolating exmtDNA from genomic DNA may also be useful in the present invention for quantifying mtDNA and exmtDNA by any of a variety of techniques well known in the art for determining the amount of specific nucleic acid target sequences present in a sample based on specific hybridization of a primer to the target sequence. Optionally, in certain of these techniques, hybridization precedes nucleotide polymerase catalyzed extension of the primer using the strand containing the target sequence as a template, and/or ligation of oligonucleotides hybridized to adjacent target sequences, and embodiments of the invention using primer extension are particularly preferred. For examples of references on such quantitative detection techniques, including those that may be used to detect nucleotide insertions, substitutions or deletions in a portion of an exmtDNA sequence site near an oligonucleotide primer target hybridization site that corresponds to a portion of the mtDNA sequence of SEQ ID NO:2, and further including those that involve primer extension, see U.S. 5,760,205 and the references cited therein, all of which are hereby incorporated by reference, and see also, for example, Botstein et al. (*Am. J. Hum. Gen.* 32:314, 1980), Gibbs et al. (*Nucl. Ac. Res.* 17:2437, 1989), Newton et al. (*Nucl. Ac. Res.*

17:2503, 1989), Grossman et al. (*Nucl. Ac. Res.* 22:4527, 1994), and Saiki et al. (*Proc. Nat. Acad. Sci.* 86:6230, 1989), all of which are hereby incorporated by reference. A particularly useful method for this purpose is the primer extension assay disclosed by Fahy et al. (*Nucl. Acids Res.* 25:3102, 1997) and by Ghosh et al. (*Am. J. Hum. Genet.* 58:325, 1996), both of which references are hereby incorporated in their entireties, as is Krook et al. (*Hum. Molec. Genet.* 1:391, 1995) which teaches modification of primer extension reactions to detect multiple nucleotide substitutions, insertions, deletions or other mutations. Other examples of useful techniques for quantifying the presence of specific nucleic acid target sequences in a sample include but need not be limited to labeled probe hybridization to the target nucleic acid sequences with or without first partially separating target nucleic acids from other nucleic acids present in the sample.

Examples of other useful techniques for determining the amount of specific nucleic acid target sequences present in a sample based on specific hybridization of a primer to the target sequence include specific amplification of target nucleic acid sequences and quantification of amplification products, including but not limited to polymerase chain reaction (PCR, Gibbs et al., *Nucl. Ac. Res.* 17:2437, 1989), transcriptional amplification systems, strand displacement amplification and self-sustained sequence replication (3SR, Ghosh et al, in Molecular Methods for Virus Detection, 1995 Academic Press, NY, pp. 287-314), the cited references for which are hereby incorporated in their entireties. Examples of other useful techniques include ligase chain reaction, single stranded conformational polymorphism analysis, Q-beta replicase assay, restriction fragment length polymorphism (RFLP, Botstein et al., *Am. J. Hum. Gen.* 32:314, 1980) analysis and cycled probe technology, as well as other suitable methods that will be known to those familiar with the art.

In a particularly preferred embodiment of the invention, primer extension is used to quantify exmtDNA and mtDNA present in a biological sample. (Ghosh et al., *Am. J. Hum. Genet.* 58:325, 1996) This embodiment may offer certain advantages by permitting both exmtDNA and mtDNA to be simultaneously quantified using a single oligonucleotide primer capable of hybridizing to a complementary nucleic acid target sequence that is present in a defined region of mtDNA and in a corresponding region of

a exmtDNA sequence. Without wishing to be bound by theory, the use of a single primer for quantification of exmtDNA and of mtDNA is believed to avoid uncertainties associated with potential disparities in the relative hybridization properties of multiple primers and may offer other advantages. Where such a target sequence is situated adjacent to an exmtDNA nucleotide sequence position that is a nucleotide substitution, insertion or deletion relative to the corresponding mtDNA sequence position, primer extension assays may be designed such that oligonucleotide extension products of primers hybridizing to mtDNA are of different lengths than oligonucleotide extension products of primers hybridizing to exmtDNA. Accordingly, the amount of exmtDNA in a sample and the amount of mtDNA in the sample may be determined by quantification of distinct extension products that are separable on the basis of sequence length or molecular mass, for purposes of calculating the ratio  $r$  as described above.

Sequence length or molecular mass of primer extension assay products may be determined using any known method for characterizing the size of nucleic acid sequences with which those skilled in the art are familiar. In a preferred embodiment, primer extension products are characterized by gel electrophoresis. In another preferred embodiment, primer extension products are characterized by mass spectrometry (MS), which may further include matrix assisted laser desorption ionization/ time of flight (MALDI-TOF) analysis or other MS techniques known to those having skill in the art. See, for example, U.S. 5,622,824, U.S. 5,605,798 and U.S. 5,547,835, all of which are hereby incorporated by reference in their entireties. In another preferred embodiment, primer extension products are characterized by liquid or gas chromatography, which may further include high performance liquid chromatography (HPLC), gas chromatography-mass spectrometry (GC-MS) or other well known chromatographic methodologies.

Any exmtDNA sequence or portion of an exmtDNA sequence that corresponds to the human mtDNA sequence of SEQ ID NO:2 or a portion thereof or several portions thereof may be useful in this embodiment of the invention. Examples of human exmtDNA sequences that are useful in this and other embodiments of the invention are disclosed in SEQ ID NO:1, in SEQ ID NO:3 and in SEQ ID NO:4.

Nucleotide positions at which exmtDNA and mtDNA differ are provided in Figure 4, in which are presented the identities of nucleotides at particular sequence positions in the human mtDNA genome (SEQ ID NO:2) and the corresponding positions in exmtDNA of SEQ ID NOS:1, 3 and 4, and in other specific exmtDNA regions corresponding to

5 the human mtDNA encoded genes NADH dehydrogenase subunit 1 (ND1), NADH dehydrogenase subunit 2 (ND2), NADH dehydrogenase subunit 3 (ND3), NADH dehydrogenase subunit 4 (ND4), NADH dehydrogenase subunit 4L (ND4L), NADH dehydrogenase subunit 5 (ND5), NADH dehydrogenase subunit 6 (ND6), tRNA<sup>Lys</sup>, tRNA<sup>Ile</sup>, tRNA<sup>Gln</sup>, tRNA<sup>Met</sup>, tRNA<sup>Cys</sup>, tRNA<sup>Tyr</sup>, tRNA<sup>Asp</sup> or cytochrome c oxidase 3

10 (CO3). These exmtDNA nucleotide sequences that correspond to the human mtDNA sequence of SEQ ID NO:2 and that differ from SEQ ID NO:2 by the specific nucleotide substitution or deletion provided in Figure 4 for an indicated nucleotide position number (using the numbering system of Anderson et al., 1981 *Nature* 290:457) were determined as provided herein by sequence analysis of SEQ ID NOS:1, 3 or of SEQ ID

15 NO:4 or of cloned PCR amplicons generated as described herein using oligonucleotide primer sets disclosed in Tables 1-3.

Portions of SEQ ID NO:2 that include nucleic acid sequences encoding the mitochondrial ETC enzymes cytochrome c oxidase 1 (CO 1), cytochrome c oxidase 2 (CO 2), ATP synthetase subunit 8 (ATPase 8) and ATP synthetase subunit 6 (ATPase 20 6) may be particularly useful, and in preferred embodiments of the invention these sequences comprise isolated nucleic acid molecules that have nucleotide sequences identical or complementary to corresponding nucleic acid sequences present in exmtDNA of SEQ ID NO:1 and/or SEQ ID NO:3 and/or SEQ ID NO:4. Portions of SEQ ID NO:2 that include nucleic acid sequences encoding the mitochondrial tRNAs, 25 including mitochondrial isoleucyl, glutamyl, methionyl, aspartyl, cysteinyl, tyrosinyl and lysyl tRNAs may also be particularly useful. Also particularly useful in preferred embodiments of the invention are sequences comprising isolated nucleic acid molecules that have nucleotide sequences identical or complementary to corresponding nucleic acid sequences present in exmtDNA of SEQ ID NO:1 and/or of SEQ ID NOS:3 and/or 30 4, and/or to corresponding regions of SEQ ID NO:2 having the specific nucleotide

substitutions or deletions identified at the indicated corresponding mtDNA nucleotide position in Figure 4, or any portion or fragment thereof.

Figure 4 shows nucleotides present in human mtDNA of SEQ ID NO:2 (Anderson et al., 1981 *Nature* 290:457) at the indicated nucleotide position number, and 5 also shows single base substitutions or deletions, which substitutions and deletions are provided by the instant invention, at the corresponding nucleotide position number in human exmtDNA sequences. The "gene" column refers to the mtDNA encoded gene region in which the indicated mtDNA nucleotide position numbers are present: "COI", "COII" and COIII" refer, respectively, to cytochrome c oxidase subunits I (COX I), II 10 (COX II) and III (COX III); "ATPase" refers to ATP synthetase (also commonly referred to as ATP synthase); "ND" refers to NADH dehydrogenase (also known as NADH:ubiquinone oxidoreductase or mitochondrial electron transport chain Complex I); and genes encoding tRNA specific for the amino acid indicated by the well known three letter code are appropriately labeled.

15 The "fragment" column in Figure 4 refers either (i) to cloned and sequenced exmtDNA containing the indicated nucleotide substitutions and deletions relative to the corresponding positions in human mtDNA, wherein "5.8 kb pg" refers to SEQ ID NO:1 and "519 bp pg" refers to SEQ ID NO:3; or alternatively, (ii) identifies oligonucleotide primer sets from Table 1 (COX I, II and III), Table 2 (ND1-4, 4L, 5 and 20 6), or Table 3 (ATP synthetase subunits 8 and 6) that were used to PCR amplify cellular DNA as described herein, using a two-digit number with a letter code (A = alternate, i.e., an alternate primer complementary to an overlapping and/or adjacent region of human mtDNA [SEQ ID NO:2] relative to a first primer having the same two-digit number and orientation -- such alternates are included to provide means for PCR 25 amplifying a particular human exmtDNA region containing one or more substitutions or deletions listed in Figure 4, where under certain conditions one member of the primer pair comprising the first primer and its alternate may be incapable of duplex formation with an exmtDNA target sequence due to an uncharacterized sequence variation, for example a mutation such as a point mutation, in the target sequence. Typically, an 30 alternate primer will be designed to anneal to an exmtDNA target sequence that

commences at a nucleotide corresponding to a nucleotide position in human mtDNA situated within 1-100 nucleotides of the position where the first primer anneals, preferably within 1-50 nucleotides, more preferably within 1-30 nucleotides and still more preferably within 1-15 nucleotides. F = forward orientation; R = reverse  
5 orientation.)

TABLE I  
OLIGONUCLEOTIDE PRIMERS CORRESPONDING TO CYTOCHROME  
C OXIDASE ENCODING MTDNA SEQUENCES

SEQ. ID. NO.	PRIMER	GENE	NUCLEOTIDE POSITION (5', light strand)	PRIMER SEQUENCE 5'->3'
14	11F	CO I	5864	GTCATAATGCTTCACTCAGCCA
15	11FA	CO I	5859	TTACAGTCCAATGCTTCACTC
16	11R	CO I	6177	TATGCGGGGAAACGCCAT
17	11RA	CO I	6180	TGTTTATGCGGGGAAACGC
18	12F	CO I	6138	GGCAACTGACTAGTCCCCTA
19	12FA	CO I	6125	AATCGGAGGCTTGCGAACTG
20	12R	CO I	6425	GTGGTATTGGGTTATGGCA
21	12RA	CO I	6422	TGGTATTGGGTTATGGCAGGG
22	13F	CO I	6383	GGCCATCAATTTCATCACAA
23	13FA	CO I	6358	TAGCAGGTGTCTCCTCTATCTT
24	13R	CO I	6697	ATACCTATGTATCCAAATGGTTCTT
25	13RA	CO I	6699	CCATACCTATGTATCCAAATGGTTC
26	14F	CO I	6657	GGAATAATCTCCCATAATTGTAACCTT
27	14R	CO I	6945	CAGGCCACCTACGGTGAA
28	14RA	CO I	6947	GTCAGGCCACCTACGGTG
29	15F	CO I	6914	AGTGCTCTGAGCCCTAGGAT
30	15FA	CO I	6902	ATGATCTGCTGCAGTGCTCT
31	15R	CO I	7193	ATTCCGGATAGGCCGAGA
32	16F	CO I	7159	TGGCGTAAATCTAACCTTCTT
33	16R	CO I	7451	GGGGTTCGATTCCCTTCCTT
34	16RA	CO I	7472	TTGGCTTGAAACCAGCTT
35	21F	CO II	7546	TTGTCAAAGTTAAATTATAGGCTA
36	21FA	CO II	7548	GTCAAAGTTAAATTATAGGCTAAA
37	21R	CO II	7832	ACCTCGTCTGTTATGTAAGGAT
38	21RA	CO II	7834	TGACCTCGTCTGTTATGTAAGG
39	22F	CO II	7792	CGCCATCATCCTAGCCTCA
40	22R	CO II	8050	ATGAGTGCAAGACGTCTTGAT
41	23F	CO II	8003	AATCGAGTAGTACTCCGATTGA

SEQ. ID. NO.	PRIMER	GENE	NUCLEOTIDE POSITION (5', light strand)	PRIMER SEQUENCE 5'->3'
42	23FA	CO II	8007	GAGTAGTACTCCGATTGAAGCC
43	23R	CO II	8286	GTTAGCTTACAGTGGGCTCTAGA
44	23RA	CO II	8287	AGTTAGCTTACAGTGGGCTCTAG
45	31F	CO III	9171	ACTTCTAGTAAGCCTCTACCTGCA
46	31FA	CO III	9173	TTCTAGTAAGCCTCTACCTGCACG
47	31R	CO III	9447	AGGTAATAAAATAGGATTATCCCGTA
48	31RA	CO III	9443	AATAAAATAGGATTATCCGTATCGA
49	32F	CO III	9416	CCACACACCACCTGTCCAA
50	32FA	CO III	9415	ACCACACACCACCTGTCCA
51	32R	CO III	9741	AAGGGAGACTCGAAGTACTCTGA
52	33F	CO III	9712	TGGGTCTCTATTTACCCCTCCTA
53	33FA	CO III	9698	TATTACAATTTACTGGGTCTCT
54	33R	CO III	10010	ACTAGTTAATTGGAAGTTAACGGTA

TABLE 2  
OLIGONUCLEOTIDE PRIMERS CORRESPONDING TO NADH  
DEHYDROGENASE ENCODING mtDNA SEQUENCES

SEQ. ID. NO.	PRIMER GENE	NUCLEOTIDE POSITION (5', light strand)	PRIMER LENGTH	PRIMER SEQUENCE 5'->3'
55	61F	ND1	3281	GAGGTTCAATTCCCTTCTTAAC
56	61R	ND1	3631	TTGAGTAAACGGCTAGGCTAGA
57	62F	ND1	3589	CTGGTCAACCTCAACCTAGG
58	62R	ND1	3946	GGCCTGCGGCGTATTC
59	63F	ND1	3908	CCGAAGGGGAGTCCGAA
60	63R	ND1	4270	ATCAAAGTAACTCTTTATCAGACA
61	71F	ND2	4447	TTGGTTATAACCCTTCCCCTA
62	71R	ND2	4769	CTATTCCTAGTTTTATTGCTATAGCT
63	72F	ND2	4699	ACAATATACTCTCCGGACAATG
64	72R	ND2	5054	GAATGGTTATGTTAGGGTTGTAC
65	73F	ND2	4990	AGCTACGCCAAAATCTTAGCATA
66	73R	ND2	5311	AGGGTGATGGTGGCTATGA
67	74F	ND2	5234	CCCGCTAACCGGCTT
68	74R	ND2	5536	AGGGCTTGAAGGCTCTT
69	81F	ND4L	10433	TTTCGACTCATTAATTATGATAAT
70	81R	ND4L	10782	CATGTCAGTGGTAGTAATATAATTG
71	82F	ND4	10718	CACATATGGCCTAGACTACGTA
72	82R	ND4	11060	ATAATTAAGGAGATTGTAGGGA
73	83F	ND4	10999	CCAACGCCACTTATCCAGT
74	83R	ND4	11342	AAGCTATTCTGTAAGCTAGTCATAT
75	84F	ND4	11275	CTCACTAACATTCTACTACTCACT
76	84R	ND4	11618	GTGGCTGATTGAAGAGTATGC
77	85F	ND4	11554	CCTATGAGGCATAATTATAACAAG
78	85R	ND4	11894	ACGTGGTTACTAGCACAGAGAG
79	86F	ND4	11834	TGACTTCTAGCAAGCCTCG
80	86R	ND4	12147	ACAATCTGATGTTTGGTAAAC
81	91F	ND6	14119	CTCATCTAACCTACTCCTAA
82	91R	ND6	14442	GCGATGGCTATTGAGGAGT
83	92F	ND6	14384	GCTAACCCCACCTAAACACT
84	92R	ND6	14693	TTCATATCATTGGTCGTGGT
85	75F	ND3	10007	TAGTACCGTTAACCTCCAATTAACT
86	75R	ND3	10430	TCATAATTAAATGAGTCGAAATCA
87	93F	NDS	12281	CAGCTATCCATTGGTCTTAGG
88	93R	NDS	12671	TATTTGAAGAACTGATTAATGTTTG
89	94F	NDS	12612	AGCATTGTTCGTTACATGGTC
90	94R	NDS	12944	GGCTTGGATTAGCGTTAGAA
91	95F	NDS	12881	TCATCCTCGCCTTAGCATGA
92	95R	NDS	13212	TTTGATGTCATTGTGTAAGG
93	96F	NDS	13156	CAAACCTAACACTATGCTTAGG
94	96R	NDS	13519	ATGTTGCGGTTCGATGAT

SEQ. ID.	PRIMER GENE	NUCLEOTIDE POSITION (5', light strand)	PRIMER LENGTH	PRIMER SEQUENCE 5'->3'
95	97F	NDS	13458	CATTGGCAGCCTAGCATTAG
96	97R	NDS	13855	GATTTTATTTAAGTTGTTGGTTAG
97	98F	NDS	13795	AAACTCACAGCCCTCGCTG
98	98R	NDS	14124	TATGTGATTAGGAGTAGGGTTAGG

TABLE 3  
OLIGONUCLEOTIDE PRIMERS CORRESPONDING TO ATP SYNTHETASE SUBUNITS 8/6  
ENCODING mtDNA SEQUENCES

SEQ. ID.	PRIMER	GENE	NUCLEOTIDE POSITION (5', light strand)	PRIMER LENGTH	PRIMER SEQUENCE 5'->3'
99	41F	ATPase 8	8292	23	GCCCACGTAAAGCTAACCTAGC
100	41R	ATPase 8	8631	22	TAGTCGGTTGTTGATGAGATAT
101	41RA	ATPase 8	8632	17	AGTCGGTTGTTGATGAG
102	42F	ATPase 6	8572	23	GGCCTACCCGCCGCAGTACTGAT
103	42R	ATPase 6	8909	23	TGTAGGTGTGCCTTGTTGGTAAGA
104	43F	ATPase 6	8866	22	ATTATAGGCTTCGCTCTAAGA
105	43FA	ATPase 6	8806	21	CCAACCACCCAACCTATCTATA
106	43R	ATPase 6	9214	22	ATATGATAGGCATGTGATTGGT
107	26F	ATPase 8	8311	25	TAGCATTAAACCTTTAAGTTAAAGA
108	26R	ATPase 8	8516	19	TCGTTCACTTTGGTTCTCA

5

As described herein, the present invention provides compositions and methods related to novel exmtDNA sequences that may differ from human mtDNA sequences at one or more nucleotide positions as disclosed in Figure 4, such as, for example, the exmtDNA sequences of SEQ ID NOS:1, 3 and 4. Details for obtaining SEQ ID NO:1 are provided below in the Examples. Those having ordinary skill in the art can also readily obtain other isolated exmtDNA sequences using well known methodologies including those provided herein and in the cited references, and further including the use of the oligonucleotide primers provided in Tables 1-3 and in the Examples. Databases (e.g., GenBank, EMBL) and methods for nucleic acid sequence analysis are also well known in the art, for example, similarity between two sequences may be readily determined using well known computer programs such as the BLAST algorithm (Altschul, *J. Mol. Biol.* 219:555-565, 1991; Henikoff and Henikoff, *Proc.*

*Natl. Acad. Sci. USA* 89:10915-10919, 1992), which is available at the NCBI website (<http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST>). Default parameters may be used. Examples of other useful computer algorithms are those used in programs such as Align and FASTA, which may be accessed, for example, at the Genestream internet website 5 of the Institut de Génétique Humaine, Montpellier, France ([www2.igh.cnrs.fr/home.eng.html](http://www2.igh.cnrs.fr/home.eng.html)) and used with default parameters.

For instance, exmtDNA sequences present in SEQ ID NOS:3 or 4 that correspond to specific regions of SEQ ID NO:2 can be readily identified, and appropriate primers selected based on the sequences provided in Tables 1-3 and the 10 exmtDNA nucleotide substitutions and deletions relative to the corresponding human mtDNA sequence (SEQ ID NO:2) provided in Figure 4. For example, SEQ ID NO:3 may be derived from, or detected in, a suitable biological sample using primer set 26 and/or primer set 41 as provided in Table 3, and methodologies essentially as provided in U.S. Patent No. 5,840,493 (and references cited therein), which is hereby 15 incorporated by reference. As another example, such an approach applied to a biological sample comprising a cloned human peripheral blood leukocyte derived genomic DNA library (Clontech, Palo Alto, CA, catalog number HL1111j) probed with a nucleotide sequence corresponding to the region of SEQ ID NO:2 encoding ATP synthase subunit 8 may provide SEQ ID NO:4, portions of which may be useful in the 20 subject invention compositions and methods. From the disclosure herein, as will be readily apparent to those having ordinary skill in the art, the nucleotide sequence from position 215-733 in SEQ ID NO:4 corresponds to SEQ ID NO:3, and both SEQ ID NOS:3 and 4 correspond to a portion of SEQ ID NO:2 that differs from SEQ ID NO:2 by having specific nucleotide substitutions disclosed in Figure 4. The nucleotide 25 sequences in SEQ ID NO:4 from positions 1-214 and 734-1263, conversely, do not correspond to mtDNA sequences of SEQ ID NO:2 and further may represent novel nucleic acid sequences. Sequence analysis (e.g., using online databases and algorithms as described above) of nucleotides 734-1263 suggests that nucleotides 963-1105 encode an exon of a human carboxypeptidase-N "L" subunit and that nucleotides 734-962 30 represent novel intron sequences within the carboxypeptidase-N gene, with nucleotides

1106-1263 representing additional novel (and probably non-coding) sequences as well. The exmtDNA sequence of nucleotides 215-733 thus may be present in human genomic DNA as a mtDNA-like pseudogene situated within intronic DNA.

In another particularly preferred embodiment of the invention, DNA in a biological sample containing exmtDNA and/or mtDNA is first amplified by methodologies well known in the art and described above, such that the amplification products may be used as templates in a method for quantifying the amount of exmtDNA and mtDNA present in the sample. Accordingly, it may be desirable to employ oligonucleotide primers that are complementary to target sequences that are identical in, and common to, mtDNA and exmtDNA, for example PCR amplification templates and primers prepared according to Fahy et al. (*Nucl. Acids Res.*, 25:3102, 1997) and Davis et al. (*Proc. Nat. Acad. Sci. USA* 94:4526, 1997; see also Hirano et al., *Proc. Nat. Acad. Sci. USA* 94:14894, 1997, and Wallace et al., *Proc. Nat. Acad. Sci. USA* 94:14900, 1997.)

Biological samples containing exmtDNA and mtDNA may comprise any tissue or cell preparation in which exmtDNA and mtDNA may be present. Biological samples may be provided by obtaining a blood sample, biopsy specimen, tissue explant, organ culture or any other tissue or cell preparation from a subject or a biological source. The subject or biological source may be a human or non-human animal, a primary cell culture or culture adapted cell line including but not limited to genetically engineered cell lines that may contain chromosomally integrated or episomal recombinant nucleic acid sequences, immortalized or immortalizable cell lines, somatic cell hybrid or cytoplasmic hybrid "cybrid" cell lines, differentiated or differentiable cell lines, transformed cell lines and the like. In certain preferred embodiments of the invention, the subject or biological source may be suspected of having or being at risk for having a disease associated with altered mitochondrial function, and in certain preferred embodiments of the invention the subject or biological source may be known to be free of a risk or presence of such as disease.

In certain other preferred embodiments where it is desirable to determine whether or not a subject or biological source falls within clinical parameters indicative

of Alzheimer's disease (AD), signs and symptoms of AD that are accepted by those skilled in the art may be used to so designate a subject or biological source, for example clinical signs referred to in McKhann et al. (*Neurology* 34:939, 1984, National Institute of Neurology, Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association Criteria of Probable AD, NINCDS-ADRDA) and references cited therein, or other means known in the art for diagnosing AD.

In certain aspects of the invention, biological samples containing mtDNA and exmtDNA may be obtained from the subject or biological source before and after contacting the subject or biological source with a candidate agent, for example 10 to identify a candidate agent capable of effecting a change in the value of the ratio  $r$ , defined above, relative to the value of  $r$  before exposure of the subject or biological source to the agent.

In a most preferred embodiment of the invention, the biological sample containing mtDNA and exmtDNA may comprise a crude buffy coat fraction of whole 15 blood, which is known in the art to comprise further a particulate fraction of whole blood enriched in white blood cells and platelets and substantially depleted of erythrocytes. Those familiar with the art will know how to prepare such a buffy coat fraction, which may be prepared by differential density sedimentation of blood components under defined conditions, including the use of density dependent separation 20 media, or by other methods.

In another most preferred embodiment of the invention, the amount of exmtDNA and mtDNA in a biological sample may be quantified by first heating the sample in water to lyse cells contained therein, and then extracting cellular DNA from the lysed cells using an aqueous DNA extraction procedure. "Heating" may involve 25 treating the cells for various times, typically 1-120 minutes, at a high temperature that is at least 80°C, preferably at least 90°C, more preferably at least 95°C and most preferably in a boiling water bath. Based on the compositions and methods disclosed in the present application, the ordinarily skilled artisan will be able to readily determine optimal times and temperatures for heating samples to practice the invention without 30 undue experimentation. As used herein, an "aqueous DNA extraction" method refers to

preparation of DNA from such a boiled cell lysate without subjecting the lysate to sodium dodecylsulfate(SDS)/proteinase K treatments and/or without fractionating the lysate using a phenol-chloroform two-phase separation extraction step. Those skilled in the art will be familiar with various standard procedures for preparing and handling

- 5 DNA without the use of SDS/ proteinase K and/or phenol-chloroform.

According to certain embodiments of the invention, the particular cell type or tissue type from which a biological sample is obtained may influence qualitative or quantitative aspects of the exmtDNA and/or mtDNA contained therein relative to exmtDNA and/or mtDNA obtained from distinct cell or tissue types of a common 10 biological source. As described above, some diseases associated with altered mitochondrial function may manifest themselves in particular cell or tissue types. For example, AD is primarily a neurodegenerative disease that particularly effects changes in the central nervous system (CNS). It is therefore within the contemplation of the invention to quantify exmtDNA and mtDNA in biological samples from different cell 15 or tissue types as may render the advantages of the invention most useful for a particular disease associated with altered mitochondrial function, and the relevant cell or tissue types will be known to those familiar with such diseases.

In order to determine whether a mitochondrial alteration may contribute to a particular disease state, it may be useful to construct a model system for diagnostic 20 tests and for screening candidate therapeutic agents in which the nuclear genetic background may be held constant while the mitochondrial genome is modified. It is known in the art to deplete mitochondrial DNA from cultured cells to produce  $\rho^0$  cells, thereby preventing expression and replication of mitochondrial genes and inactivating mitochondrial function. See, for example, International Publication Number WO 25 95/26973, which is hereby incorporated by reference in its entirety, and references cited therein. It is further known in the art to repopulate such  $\rho^0$  cells with mitochondria derived from foreign cells in order to assess the contribution of the donor mitochondrial genotype to the respiratory phenotype of the recipient cells. Such cytoplasmic hybrid cells, containing genomic and mitochondrial DNAs of differing biological origins, are 30 known as cybrids.

“ $\rho^0$  cells” are cells essentially completely depleted of mtDNA, and therefore have no functional mitochondrial respiration/ electron transport activity. Such absence of mitochondrial respiration may be established by demonstrating a lack of oxygen consumption by intact cells in the absence of glucose, and/or by demonstrating 5 a lack of catalytic activity of electron transport chain enzyme complexes having subunits encoded by mtDNA, using methods well known in the art. (See, e.g., Miller et al., *J. Neurochem.* 67:1897-1907, 1996.) That cells have become  $\rho^0$  cells may be further established by demonstrating that no mtDNA sequences are detectable within the cells. For example, using standard techniques well known to those familiar with the 10 art, cellular mtDNA content may be measured using slot blot analysis of 1  $\mu$ g total cellular DNA probed with a mtDNA-specific oligonucleotide probe radiolabeled with, e.g.,  $^{32}\text{P}$  to a specific activity  $\geq$  900 Ci/gm. Under these conditions  $\rho^0$  cells yield no detectable hybridizing probe signal. Alternatively, any other method known in the art for detecting the presence of mtDNA in a sample may be used that provides comparable 15 sensitivity.

“Mitochondrial DNA depleted” cells (“mtDNA depleted cells”) are cells substantially but not completely depleted of functional mitochondria and/or mitochondrial DNA, by any method useful for this purpose. MtDNA depleted cells are preferably at least 80% depleted of mtDNA as measured using the slot blot assay 20 described above for the determination of the presence of  $\rho^0$  cells, and more preferably at least 90% depleted of mtDNA. Most preferably, mtDNA depleted cells are depleted of >95% of their mtDNA.

Mitochondria to be transferred to construct model systems in accordance with the present invention may be isolated from virtually any tissue or cell source. Cell 25 cultures of all types may potentially be used, as may cells from any tissue. However, fibroblasts, brain tissue, myoblasts and platelets are preferred sources of donor mitochondria. Platelets are the most preferred, in part because of their ready abundance, and their lack of nuclear DNA. This preference is not meant to constitute a limitation on the range of cell types that may be used as donor sources.

For example, platelets may be isolated by an adaptation of the method of Chomyn (*Am. J. Hum. Genet.* 54:966-974, 1994). However, it is not necessary that this particular method be used. Other methods are easily substituted. For instance, if nucleated cells are used, cell enucleation and isolation of mitochondria isolation can be 5 performed as described by Chomyn et al., *Mol. Cell. Biol.* 11:2236-2244, 1991. Human tissue from a subject suspected of having or being at risk for having a disease associated with altered mitochondrial function, or from a subject known to be free of a risk or presence of such a disease, may be the source of donor mitochondrial DNA.

After preparation of mitochondria by isolation of platelets or enucleation 10 of donor cells, the mitochondria may be transplanted into  $\rho^0$  cells or mtDNA depleted cells using any known technique for introducing an organelle into a recipient cell, including but not limited to polyethylene glycol (PEG) mediated cell membrane fusion, cell membrane permeabilization, cell-cytoplasm fusion, virus mediated membrane fusion, liposome mediated fusion, particle mediated cellular uptake, microinjection or 15 other methods known in the art. For example by way of illustration and not limitation, mitochondria donor cells ( $\sim 1 \times 10^7$ ) are suspended in calcium-free Dulbecco's modified Eagle (DME) medium and mixed with  $\rho^0$  cells ( $\sim 0.5 \times 10^6$ ) in a total volume of 2 ml for 5 minutes at room temperature. The cell mixture is pelleted by centrifugation and resuspended in 150  $\mu$ l PEG (PEG 1000, J.T. Baker, Inc., 50% w/v in DME). After 1.5 20 minutes, the cell suspension is diluted with normal  $\rho^0$  cell medium containing pyruvate, uridine and glucose, and maintained in tissue culture plates. Medium is replenished daily, and after one week medium lacking pyruvate and uridine is used to inhibit growth of unfused  $\rho^0$  cells. These or other methods known in the art may be employed to produce cytoplasmic hybrid, or "cybrid", cell lines.

As a non-limiting example, cybrid model systems may be useful for 25 diagnosing a patient suspected of having or being at risk for a disease associated with altered mitochondrial function. According to this example, the patient's mitochondria are used to construct cybrid cells as described above. These cybrid cells may then be propagated in vitro and used to provide a biological sample for the determination of the 30 ratio  $r$ , which can be compared to an  $r$  value calculated from samples of a control cybrid

cell line constructed with mitochondria from a subject known to be free of disease. Where it may be desirable to compare the influence upon  $r$  of mitochondria from different sources, both cybrid cell lines may be constructed from the same  $\rho^0$  cell line to provide a constant background environment. These and similar uses of model systems  
5 according to the invention for determining the risk for or presence of a disease associated with altered mitochondrial function will be appreciated by those familiar with the art and are within the scope and spirit of the invention.

As another non-limiting example, cybrid model systems may be useful for identifying agents suitable for treating a disease associated with altered  
10 mitochondrial function. According to this example, a cybrid cell line may be a biological source in which the ratio  $r$  is calculated as described above, before and after cybrid cells are contacted with a candidate agent for treating disease. Such a cybrid cell line may be used to screen candidate agents by identifying those agents capable of effecting a change in the value of  $r$  relative to the value of  $r$  before exposure to the  
15 agent. The present invention thus provides model systems for selecting therapeutic agents that may be suitable for the treatment of diseases associated with altered mitochondrial function. These and similar uses of model systems according to the invention for the screening and identification of agents that influence the ratio  $r$  defined above, will be appreciated by those familiar with the art and are within the scope and  
20 spirit of the invention.

In addition, although the present invention is directed primarily towards model systems for diseases in which the mitochondria have metabolic alterations, it is not so limited. Conceivably there are disorders wherein mitochondria contain structural or morphological defects or anomalies, and the model systems of the present invention  
25 are of value, for example, to find drugs that can address that particular aspect of the disease. Also, there are certain individuals that have or are suspected of having extraordinarily effective or efficient mitochondrial function, and the model systems of the present invention may be of value in studying such mitochondria. Moreover, it may be desirable to put known normal mitochondria into cell lines having disease  
30 characteristics, in order to rule out the possibility that mitochondrial alterations

contribute to pathogenesis. All of these and similar uses are within the scope of the present invention, and the use of the phrase "mitochondrial alteration" herein should not be construed to exclude such embodiments.

According to the present invention, a ratio  $r$  as defined herein is determined in a biological sample, for example by calculation following quantification of mtDNA and exmtDNA using a technique based on specific oligonucleotide hybridization to a target sequence. This hybridization may be optionally followed by target template directed extension, such as in primer extension assays described herein. For certain diseases associated with altered mitochondrial function, calculation of  $r$  may have diagnostic usefulness. For example, where other clinical indicators of a disease associated with altered mitochondrial function are known, values for  $r$  in subjects known to be free of a risk or presence of such disease based on the absence of these indicators may be determined to establish a control range for  $r$ . The ratio may also be calculated in biological samples obtained from subjects suspected of having or being at risk for having a disease associated with altered mitochondrial function, and compared to the control range of  $r$  values determined in disease free subjects. Those having familiarity with the art will appreciate that there may be any number of variations on the particular subjects, biological sources and bases for comparing  $r$  values that are useful beyond those that are expressly presented herein, and these additional uses are within the scope and spirit of the invention.

For instance, determination of  $r$  in may take the form of a diagnostic assay performed on whole blood collected from a subject by routine venous blood draw, on buffy coat cells prepared from blood or on biological samples that are other cells, organs or tissue from a subject. Alternatively, in certain situations it may be desirable to construct cybrid cell lines using mitochondria from either control subjects or subjects suspected of being at risk for a disease associated with altered mitochondrial function. Such cybrids may be used to determine  $r$  for diagnostic purposes, or as biological sources for screening assays to identify agents that may be suitable for treating disease based on their ability to change the  $r$  value obtained from treated cells. In one embodiment of this aspect of the invention, therapeutic agents or combinations of

agents that are tailored to effectively treat an individual patient's particular disease may be identified by routine screening of candidate agents on cybrid cells constructed with the patient's mitochondria.

The present invention provides compositions and methods that are useful  
5 in pharmacogenomics, for the classification and/or stratification of a subject or a patient population, for instance correlation of one or more traits in a subject with indicators of the responsiveness to, or efficacy of, a particular therapeutic treatment. In one aspect of the invention, measurement of  $r$  in a biological sample from a subject is combined with identification of the subject's apolipoprotein E (APOE) genotype to determine the risk  
10 for, or presence of, Alzheimer's disease (AD) in the subject. The apolipoprotein E type 4 allele (*APOE-ε4*) allele is a genetic susceptibility factor for sporadic AD and confers a two fold risk for AD (Corder et al., *Science* 261:921, 1993; *see also* "National Institute on Aging/Alzheimer's Association Working Group Consensus Statement," *Lancet* 347:1091, 1996 and references cited therein, all of which are hereby incorporated by  
15 reference in their entireties.). Accordingly, in a preferred embodiment of the invention, the method for determining the risk for or presence of AD in a subject by comparing  $r$  values will further comprise determining the APOE genotype of the subject suspected of being at risk for AD. By using the combination of the methods for determining  $r$ , as disclosed herein, and methods known in the art for determining APOE genotype, an  
20 enhanced ability to detect the relative risk for AD is provided by the instant invention along with other related advantages. Similarly, where APOE genotype and risk for AD are correlated, the present invention provides advantageous methods for identifying agents suitable for treating AD where such agents affect  $r$  in a biological source.

As described herein, determination of  $r$  may be used to stratify an AD  
25 patient population. Accordingly, in another preferred embodiment of the invention, determination of  $r$  in a biological sample from an AD subject may provide a useful correlative indicator for that subject. An AD subject so classified on the basis of an  $r$  value may then be monitored using AD clinical parameters referred to above, such that correlation between  $r$  value and any particular clinical score used to evaluate AD may  
30 be monitored. For example, stratification of an AD patient population according to  $r$

values may provide a useful marker with which to correlate the efficacy of any candidate therapeutic agent being used in AD subjects. In a further preferred embodiment of this aspect of the invention, determination of  $r$  in concert with determination of an AD subject's APOE genotype may also be useful. These and 5 related advantages will be appreciated by those familiar with the art.

In another aspect, the invention provides exmtDNA sequences that may be useful in the detection of telomeric events that are related to diseases, including diseases associated with altered mitochondrial function, or in the identification of agents that are suitable for the treatment of such diseases. Telomeres are specialized DNA sequences 10 found at the ends of eukaryotic chromosomes that are complexed *in vivo* with specific telomere-associated proteins. Typically, telomeres comprise tandemly repeated, highly conserved G-rich sequences (for example, the human repeat sequence is 5'-TTAGGG). The eukaryotic DNA polymerase  $\alpha$  is unable to replicate the very ends of linear DNA and thus every replication cycle results in a shortening of the telomeric ends in normal 15 somatic cells. Without wishing to be bound by theory, this progressive shortening of telomeric length may contribute to molecular events that lead to cellular senescence (e.g., a "mitotic clock"; see Harley et al., *Nature* 345:458-460, 1990; Greider, *Bioessays* 12:363-369, 1990, hereby incorporated by reference). According to this model, in order to avoid telomeric shortening germline cells employ the enzyme telomerase, a 20 ribonucleoprotein that catalyzes the addition of the tandem repeat sequence to the ends of chromosomes (Kim et al., *Science* 226:2011-2015, 1994). Telomerase is also expressed in some tumors, suggesting that telomerase activation may be one means by which tumor cells develop from somatic cells (Shay et al., *Curr. Opin. Oncol.* 8:66-71, 1996).

25 Dynamic processes in the telomeric or distal end regions of chromosomes that involve specific nucleic acid sequences, and in particular that may involve particular nucleotide polymerase and nuclease activities, have thus been implicated in chromosomal events that may be related to cellular and molecular mechanisms of disease. See also, for example, Fossel, *J. Amer. Med. Assoc.* 279:1732 30 (1998); LaBranche et al., *Nat. Genet.* 19:199 (1998); Shay, *Cancer J. Sci. Am.* 4:526

(1998); Nowak et al., *Cancer J. Sci. Am.* 4:148 (1998); Iwama et al., *Hum. Genet.* 102:397 (1998), all of which are hereby incorporated by reference. In one embodiment of the invention, nucleic acid sequences are provided that may be used to monitor telomeric events, including but not limited to telomerase activity. Accordingly, because 5 telomeric structure may be related to cellular growth potential and/or senescence, as noted above by way of non-limiting theory, nucleic acid based determination of telomeric structure may provide effective means for the detection of related disease processes. The present invention provides identification of human exmtDNA sequences in human chromosomal telomeric regions, and other related advantages.

10 The following examples are offered by way of illustration, and not by way of limitation.

### EXAMPLES

#### EXAMPLE 1

##### CLONING OF EXMTDNA FROM PERIPHERAL 15 BLOOD LYMPHOCYTE GENOMIC DNA LIBRARY

Peripheral blood lymphocytes were separated from the peripheral blood mononuclear cell fraction of freshly drawn venous blood from healthy human volunteers and the DNA extracted by standard techniques. Plasmid isolation, production of competent cells, transformation and manipulations using cloning vectors 20 were performed essentially as described (Sambrook et al., *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). The human lymphocyte DNA was partially digested with HindIII and inserted into the pBeloBAC11 vector (Genome Systems, Inc., St. Louis, MO with insert sizes ranging from 50 kb to over 240 kb to create a human genomic library. This library was 25 screened by PCR using the following primers, which are complementary to mtDNA sequences in the CO1 encoding region of SEQ ID NO:2 but have single-base substitutions at the 3' end. Details of PCR reaction conditions are provided in US 5,565,323, which is hereby incorporated by reference, using 30 cycles instead of 25

cycles at 95°C and with a single reaction being performed instead of five separate reactions.

5' -CCTTACACCTAGCAGGTA                   SEQ ID NO:5

5' -ACGCCGATGAATATGATAGC               SEQ ID NO:6

5                   A single positive clone was identified having a genomic DNA insert that included exmtDNA and was expanded, with the DNA then being purified using Magnum KB-100 columns (Genome Systems, St. Louis, MO). Portions of the insert were amplified with internal PCR primers complementary to human mtDNA sequences (Anderson et al., *Nature* 290:456, 1981) and sequenced using Prism DyeDeoxy  
10 terminator chemistry (Perkin-Elmer, Foster City, CA) according to the manufacturer's instructions. Sequence information at the junctions of exmtDNA and adjoining non-  
mtDNA in the positive pBeloBAC clone was obtained using the BigDye Terminator  
cycle sequencing kit (Perkin-Elmer) according to the supplier's recommendations. The  
products of sequencing reactions were purified by ethanol precipitation or by using  
15 CentriSep spin columns (Princeton Separations, Princeton, NJ), then electrophoretically  
separated using an Applied Biosystems Model 373A DNA sequencing system (Applied  
Biosystems Division of Perkin-Elmer, Foster City, CA). Sequence Navigator software  
(Applied Biosystems) was used to analyze exmtDNA sequence data, and nucleotide  
insertions, deletions or substitutions were identified by comparing exmtDNA sequences  
20 to published human mtDNA sequence data. (SEQ ID NO:2, Anderson et al., *Nature*  
290:456, 1981).

The insert contained the 5,840 base pair contiguous exmtDNA sequence of SEQ ID NO:1 (nucleotide positions 445-6284 in Figure 1), which corresponded to and exhibited 98% sequence homology with nucleotide positions 3914-9755 of the  
25 human mtDNA sequence of SEQ ID NO:2. There were 89 nucleotide positions at which substitutions or deletions (2) were detected, as indicated in Figure 4. The complete ~5.8 kb exmtDNA sequence is in a single reading frame relative to the corresponding region of the human mitochondrial genome (SEQ ID NO:2) with the exception of a two base pair deletion at nucleotide positions 8196-8197. Starting at the  
30 5' terminus (nt 445 in Fig. 1) and proceeding in the 3' direction, the exmtDNA

sequence includes DNA sequences corresponding, in order, to a truncated ND1 gene, complete ND2 and CO1 genes, a CO2 gene with the above noted two base pair deletion, complete ATP synthetase subunit 8 and ATP synthetase subunit 6 genes, interspersed tRNA genes (as indicated in Figure 4) and a truncated CO3 gene of the published 5 human mtDNA sequence of SEQ ID NO:2 (Anderson et al., *Nature* 290:456, 1981). The non-mitochondrial DNA sequences on either side of the 5,840 base pair exmtDNA sequence (nucleotides 1-444 and 6285-6691) did not display homology to any nuclear DNA sequences listed in the GenBank database.

EXAMPLE 2  
DETECTION OF EXMTDNA IN RHO-0 CELLS

In order to verify that presumptive exmtDNA sequences originated from nuclear and not mitochondrial DNA present in the DNA preparation from which the 5 human genomic library was constructed, two established cell lines were depleted of mtDNA using ethidium bromide to generate ρ<sup>0</sup> cells (Miller et al., *J. Neurochem.* 67:1897, 1996) and assayed for the presence of exmtDNA sequences. Briefly, p0118/5 and 064/5 SH-SY5Y neuroblastoma cells and 0A431 epidermal carcinoma cells were produced and maintained as described (Miller et al., 1996). Cells were harvested and 10 DNA was extracted with DNAzol (Molecular Research Center, Inc., Cincinnati, OH) according to the manufacturer's instructions. The recovered DNA was amplified by PCR and analyzed by primer extension assay using primers specific for a sequence within the mitochondrial COXI gene region containing the nucleotide at position 7146 of the mtDNA sequence (SEQ ID NO:2), which corresponds to a nucleotide substitution 15 in the exmtDNA sequence. (Figure 4 and SEQ ID NO:1) PCR primers and reaction conditions and primer extension assays were as described in Fahy et al. (*Nucl. Acids Res.* 25:3102, 1997), which is hereby incorporated by reference in its entirety. Primer extension products corresponding to a region of the exmtDNA sequence of SEQ ID NO:1 and including a nucleotide corresponding to the guanosine residue at position 20 7146 or its complement were detected in ρ<sup>0</sup> cells that contained no detectable mtDNA as described herein and in Miller et al. (*J. Neurochem.* 67:1897, 1996).

## EXAMPLE 3

## ABSENCE OF DETECTABLE TRANSCRIPTS OF EXMTDNA SEQUENCES IN RT-PCR ASSAY

The reverse transcription-polymerase chain reaction (RT-PCR; Rappolee et al., *Science* 241:708, 1991; Chelly et al., *Nature* 333:858, 1988; Brenner et al., 5 *BioTechniques* 7:1096, 1989) technique was employed to determine whether the novel 5.8 kb exmtDNA sequence cloned from a human genomic DNA library is transcribed. The buffy coat fraction of freshly drawn human venous blood was prepared using Accuspin devices (Sigma, St. Louis, MO) according to the manufacturer's instructions and total RNA was extracted from isolated buffy coat cells with Trizol reagent 10 (GibcoBRL, Bethesda, MD) as recommended by the supplier. First strand cDNA was synthesized from poly-A+ mRNA using the SuperScript™ preamplification system (GibcoBRL) with oligo(dT) as primer according to the manufacturer's instructions. PCR was conducted using this cDNA as template and primers complementary to portions of the mitochondrial cytochrome c oxidase subunit 2 (COII, COX2) mtDNA 15 sequence (SEQ ID NO:2) as described in Fahy et al. (*Nucl. Ac. Res.* 25:3102, 1997). The corresponding region of exmtDNA (SEQ ID NO:1) contains nucleotide substitutions at positions 7650 and 7868, relative to mtDNA (Figure 4). Amplicons were purified and analyzed by the primer extension assay as described in Fahy et al. such that readily distinguishable products are predicted depending on whether or not the 20 primer has hybridized to a target sequence adjacent to a sequence having the substitutions. Based on quantitative analysis of fluorescent band intensities of the primer extension products, mRNA encoding human COII gene products of mtDNA (SEQ ID NO:2) was detectable, but no mRNA encoding products from the corresponding exmtDNA region of SEQ ID NO:1 was detected, indicating that the 25 exmtDNA sequence is not expressed.

## EXAMPLE 4

## PRIMER EXTENSION ASSAY TO QUANTIFY EXMTDNA AND MTDNA

A competitive primer extension approach was used to simultaneously detect both mtDNA and exmtDNA sequences with a fluorophor-labeled primer and a selected mix of deoxynucleotides (dNTPs) and dideoxynucleotides (ddNTPs). The exmtDNA and mtDNA compete as templates in the primer extension reaction and are distinguished by differential extension of the primer. Thus, to determine the relative quantities of defined portions of mtDNA (SEQ ID NO:2) and exmtDNA (SEQ ID NO:1) using the nucleotide substitution at position 7650 (Figure 4) with the nucleotide combination of dATP, dTTP and ddGTP, the primer is extended by one base when the template is mtDNA. When exmtDNA is present as template, the primer is extended by three bases. The proportion of exmtDNA in relation to mtDNA is estimated by comparing the ratio of fluorescence intensities of the gel-separated extension products with a standard curve generated from known mtDNA/exmtDNA mixtures. (Fahy et al., 1997)

Thermo Sequenase™, dNTPs and ddNTPs were purchased from Amersham (Cleveland, OH). Calf intestine alkaline phosphatase and biochemical reagents were obtained from Boehringer Mannheim (Indianapolis, IN) and QIAquick PCR purification kits from Qiagen (Chatsworth, CA). Accuspin™ Tubes and HISTOPAQUE® 1077 were purchased from Sigma (St. Louis, MO) and EDTA vacutainers from Beckton Dickinson (San Jose, CA). *Ultma*™ DNA polymerase, *AmpliTaq*® DNA polymerase and reagents for DNA synthesis were purchased from Perkin Elmer (Foster City, CA). Oligonucleotides were synthesized on an Applied Biosystems 394 DNA/RNA synthesizer (Perkin Elmer) using standard phosphoramidite chemistry. 5' Fluorescein-labeled oligonucleotide primers were obtained by using the 6-FAM Amidite reagent in the last step of the automated synthesis. Tritylated and fluorescein-labeled oligonucleotides were purified by reverse phase chromatography using an acetonitrile gradient in 0.1 M triethylammonium acetate, pH 6.8 running buffer. The oligonucleotides migrated as single bands on a 15% denaturing polyacrylamide gel. The homogeneity of the fluorescein-labeled oligonucleotides was

independently assessed by electrophoresis on an Applied Biosystems Model 373 Sequencing System.

After IRB approval and informed consent, fresh venous blood samples were drawn from 837 patients with clinical diagnosis of probable Alzheimer's disease (AD mean age =  $74.7 \pm 1.1$  years) and 191 controls ( $67 \pm 1$  years; cognitively normal age-matched N = 114; cortico-basal ganglionic degeneration N = 2; Pick's disease N = 1; Parkinson's disease N = 24; non-insulin dependent diabetes mellitus N = 29; insulin dependent diabetes mellitus N = 6; Leigh's syndrome, N = 2, Machado Joseph Disease N = 2; idiopathic renal glycosuria N = 1; progressive supranuclear palsy N = 1; sporadic amyotrophic lateral sclerosis N = 6; familial sporadic amyotrophic lateral sclerosis N = 1; familial Alzheimer's disease N = 2). AD patients met the National Institute of Neurological, Communicative Disorders and Stroke and Alzheimer's Diseases and Related Disorders Association (NINCDS-ADRDA) criteria of probable Alzheimer's disease (McKhann et al., *Neurology* 34:939, 1984).

Blood samples were collected in EDTA vacutainers and kept at  $4^{\circ}\text{C}$  for no more than 24 h. The platelet/white blood cell fraction was isolated with Accuspin<sup>TM</sup> Tubes (Sigma Diagnostics, St. Louis, MO) using the following procedure. Three ml of HISTOPAQUE<sup>®</sup> 1077 (Sigma) were added to the upper chamber of each Accuspin<sup>TM</sup> Tube and the device was centrifuged at 1000 g for 30 s. Two to three ml of blood were then introduced into the upper chamber and separated by centrifugation at 1000 g for 10 min at room temperature. After centrifugation, the plasma and white blood cell layers were transferred to a new tube and the white blood cells were sedimented by centrifugation at 7,000 g for 10 min. The white cell pellet was resuspended in 0.4 ml of a solution containing 0.9% sodium chloride/1mM EDTA and stored at  $-80^{\circ}\text{C}$  until use.

Frozen white blood cells (0.2 mL) were thawed and were sedimented by centrifugation at 12,000 g for 5 min. The white cell pellet was washed twice with 0.6 ml of Dulbecco's Phosphate Buffered Saline (PBS; GibcoBRL, Bethesda, MD) and resuspended in 0.2 ml of water. The cells were lysed by incubation in a boiling water bath for 10 min. After cooling to room temperature, the cellular debris was sedimented by centrifugation at 14,000 g for 2 min. The supernatant was transferred to a new vial

and the approximate concentration of the crude DNA preparation was estimated from its A<sub>260</sub> absorbance. The DNA sample was stored at -80°C.

Primer extension reaction templates were prepared by PCR amplification of cellular DNA. The reactions were carried out in a total volume of 50 µl using the 5 primer pair sets described below. Following amplification, the PCR products were analyzed by electrophoresis on a 0.8% agarose gel. Reactions for analyzing exmtDNA to mtDNA ratios contained ~1 µg of cellular DNA, 2.5 U of *AmpliTaq*® DNA polymerase, 20 pmol each of the light strand primer

10 5'-CATGCAGCGCAAGTAGGTCTACAAGAC-3' [SEQ ID NO:7]

and the heavy strand primer

5'-TGTTATGTAAAGGATGCGTAGGGATGG-3' [SEQ ID NO:8]

15 and 10 nmol of each dNTP in PCR buffer (10 mM Tris. HCl, pH 8.3, 50 mM KCl, 2 mM MgCl<sub>2</sub>). After an initial denaturation step at 95°C for 10 s in a Gene Amp PCR System 9600 thermal cycler (Perkin Elmer), amplification was carried out for 30 cycles under the following conditions: 95°C for 1 min, 60°C for 1 min, 72°C for 1 min. After 20 the last cycle, reaction tubes were kept at 72°C for 4 min to ensure extension of incomplete strands to the full length 255 base pair size product.

Residual nucleotides that persisted after the PCR reaction were dephosphorylated by adding 1 unit of calf intestine alkaline phosphatase (CAP) in 5 µL of 10X CAP buffer (100 mM Tris. HCl, pH 8.3, 10 mM MgCl<sub>2</sub>, 10 mM ZnCl<sub>2</sub>) to the 25 PCR reaction mixture and incubating for 30 min at 37°C in thermal cycler. Then 1.1 µL of 0.25 M EDTA, pH 8.0 was added and the alkaline phosphatase was denatured at 75°C for 10 min.

Double-stranded PCR products were separated from primers, nucleosides and enzymes using QIAquick™ columns (Qiagen, Chatsworth, CA) and the buffers 30 provided by the manufacturer. Thus, 250 µl of buffer PB were added to the PCR

reaction mixture and mixed. A Qiaquick™ spin column (Qiagen) was placed in a 2 ml collection tube and the sample was loaded. The sample was centrifuged for 30-60 s at 14000 g and the flow through was discarded. The adsorbed PCR product was washed with 750 µl of buffer PE, and eluted with 50 µl of 10 mM Tris. HCl, pH 8.5. The 5 purified product solution was dried in a Savant SpeedVac Concentrator and then reconstituted in 20 µl of water.

The fluorescein-labeled primer for analysis of the AD-associated mutation at nucleotide position 7650 was

10

TATGAGGGCGTGATCATGAAAG

[SEQ ID NO:9]

using dATP and dTTP plus dideoxyGTP (ddGTP) to generate primer extension products of 23 and 25 nucleotides in length from mtDNA and exmtDNA templates, respectively.

15

Stock solutions of each dNTP and ddNTP were prepared by mixing equimolar amounts of the nucleotides with MgCl<sub>2</sub> and diluting the mixture to the desired concentration with 10 mM Tris, 1 mM EDTA, pH 8.0 (TE). The fluorescein-labeled primers were diluted in TE to provide final stock concentrations of 40 fmol/µl. One µl of the purified PCR-amplified DNA fragment was used as template for each 20 assay.

20

Primer extension reactions were performed in a total volume of 8 µL. The thermostable enzyme, *U1Tma*™ DNA polymerase (Perkin-Elmer, Foster City, CA) was used in primer extension assays for analyzing exmtDNA to mtDNA ratios. The reactions contained template, 20 fmol fluorescein-labeled primer, 400 µM ddNTPs/25 µM dNTPs of the appropriate nucleotide combination and 0.6 unit of enzyme in buffer containing 10 mM Tris-HCl, pH 8.8, 10 mM KC1, 0.002% Tween 20, 2 mM MgCl<sub>2</sub>. Each set of primer extension assays included control reactions with mtDNA and exmtDNA templates.

25

The products of the primer extension reaction were analyzed on an ABI 30 373 Sequencer using a 12% denaturing polyacrylamide gel and Tris borate/EDTA as

running buffer. Prior to electrophoresis, the samples in loading dye were denatured for 3 min at 85°C. Three  $\mu$ l aliquots of the standards (primer with no added template, reaction products from control DNA templates) and each unknown reaction mixture were then loaded and electrophoresed according the manufacturer's instructions.

5 Fluorescent band intensities associated with the primer extension products were estimated by the GENESCAN™ 672 software program (Perkin Elmer, Applied Biosystems Division). Quantitative analysis was carried out by correlating the fluorescent band intensities of mtDNA and exmtDNA-derived extension products from unknown samples with those from mtDNA and exmtDNA control templates.

**EXAMPLE 5**  
**PRIMER EXTENSION ASSAY USING DNA**  
**ISOLATED FROM AFFINITY ISOLATED MITOCHONDRIA**

Mitochondria were immunoaffinity isolated from cultured cells and then  
5 DNA was extracted from the isolated organelles, to determine whether detectable  
exmtDNA was present. Normal SH-SY5Y neuroblastoma cells, and normal and  $\rho^0$   
A431 epidermal carcinoma cells were produced and maintained as described (Miller  
et al., *J. Neurochem.* 67:1897, 1996). Cells were harvested by scraping in MSB (0.21  
M mannitol, 0.07 M sucrose, 0.05 M Tris-HCl, 0.01 M EDTA, pH 7.4;  $10^7$  cells/ml)  
10 and lysed by three freeze-thaw cycles. Cellular debris was removed by centrifugation at  
1000 x g for 5 min. The mitochondria enriched supernatants were used for subsequent  
immunopurification steps. The buffy coat fraction of whole blood containing white  
blood cells and platelets as described above was also prepared and lysed by freeze-thaw  
as was done with the cultured cells, to produce a mitochondrial fraction for  
15 immunopurification.

Monoclonal anti-mitochondrial antibody MAB 1273 (Chemicon  
International, Temecula, CA) was added to the mitochondria enriched fractions for 2 hr.  
Antibody-mitochondria complexes were isolated using a secondary antibody bound to  
magnetic beads (Dynal Inc., Great Neck, NY) according to the manufacturer's  
instructions. After extensive washing of the bead-bound antibody-mitochondria  
20 complexes with PBS/0.1% BSA, mtDNA was extracted from the complexes using  
DNAzol reagent (Molecular Research Center, Inc., Cincinnati, OH) according to the  
manufacturer's recommendations. Quantification of mtDNA and exmtDNA was  
performed using the competitive primer extension assay described above in Example 4.  
25 No DNA could be amplified from the DNA extracted from immunopurified  
mitochondria from the  $\rho^0$  A431 cell line, consistent with the depletion of mtDNA that is  
characteristic of the  $\rho^0$  state. Primer extension results indicated that no exmtDNA  
sequences were present in the DNA extracted from mitochondria of blood buffy coat  
cells, SH-SY5Y cells or A431 cells, confirming that exmtDNA sequences are of  
30 extramitochondrial origin.

## EXAMPLE 6

## DETERMINATION OF APOE GENOTYPE BY PRIMER EXTENSION ASSAY

Primer extension assay procedures, essentially as described above in Example 4 but with the modifications indicated here, were used to determine APOE genotype in a panel of 837 clinically diagnosed AD individuals and 191 control subjects (cognitively normal age matched individuals, patients with non-insulin dependent diabetes mellitus (NIDDM) and neurological controls). The APOE allele distribution in the AD population and controls in the study is shown in Table 4.

10 TABLE 4  
APOE GENOTYPE DISTRIBUTION

APOE allele	Control (N = 191)	AD (N = 837)
2/2	1	
2/3	16	33
2/4	3	22
3/3	112	295
3/4	56	403
4/4	3	84

DNA isolation and preparation of primer extension reaction templates by PCR were essentially as described above in Example 4, except that for *APOE* analysis, a modification of the protocol of Livak and Hainer (1994) was used. Thus, for template preparation by PCR each reaction contained ~1 µg of cellular DNA, 2.5 U of 15 *AmpliTaq*® DNA polymerase, 20 pmol each of the forward primer

5'-GGCACGGCTGTCCAAGG-3'

[SEQ ID NO:10]

20 and the reverse primer

5'-CCCGGCCTGGTACACTG-3'

[SEQ ID NO:11]

and 10 nmol of each dNTP in PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub> supplemented with 5% DMSO). After an initial denaturation step at 95°C  
 5 for 10 s in a Gene Amp PCR System 9600 thermal cycler (Perkin Elmer), amplification was carried out for 25 cycles under the following conditions: 95°C for 1 min, 55°C for 1 min, 72°C for 1 min. The PCR product obtained after amplification was 226 base pairs in length.

The fluorescein-labeled primers and nucleotide combinations for primer  
 10 extension analysis of Apo-E alleles are shown in Table 5.

TABLE 5

Gene	Primer Sequence (5'->3')	Size	Primer Extension Product Sizes	dNTP	ddNTP
APOE Codon 112	GCGGACATGGAGGACGTG SEQ ID NO:12	18	19, 20	T	G, C
APOE Codon 158	CGATGCCGATGACCTGCAGAAG SEQ ID NO:13	22	23, 24	T	G, C

15 Thermo Sequenase™ (Amersham, Cleveland, OH) DNA polymerase catalyzed reactions for Apo-E allele analysis were performed with 20 fmol fluorescein-labeled primer, 25 μM each of the appropriate ddNTP/dNTP combination and 0.64 unit of enzyme in buffer containing 10 mM Tris-HCl, pH 9.5, 5 mM KC1, 0.002% Tween 20, 2 mM MgCl<sub>2</sub>. Each set of primer extension assays included Apo-E allelic DNA controls. After an initial denaturation step at 95°C for 2 min, the reaction conditions comprised 20 cycles of 95°C for 20 s and 55°C for 40 s. The samples were concentrated to ~1 μl by heating open reaction tubes at 94°C for 7 min. After the concentration step, 8 μl of loading dye (0.5% blue dextran in 83% formamide/8.3 mM EDTA, pH 8.0) was added.

25 The products of the primer extension reaction were analyzed as described above in Example 4. The Apo-E allele composition of unknown DNA

samples was deduced by comparing the electrophoretic pattern of primer extension products with those from Apo-E allele DNA standards. APOE genotype distribution is shown in Table 4.

#### EXAMPLE 7

5

#### CORRELATION OF EXMTDNA:MTDNA RATIO WITH RISK FOR AD

Primer extension assays as described in Example 4 were used to quantify the amounts of mtDNA and exmtDNA in blood samples from a panel of 837 clinically diagnosed AD individuals and 191 control subjects (cognitively normal age matched individuals, patients with non-insulin dependent diabetes mellitus (NIDDM) and 10 neurological controls). For each subject, a ratio  $r$  was calculated using the formula:

$$r = x/(x + y)$$

wherein

15

$x$  is the amount of exmtDNA in a sample, and

$y$  is the amount of mtDNA in the sample.

20

The values for  $r$  were multiplied by 100 to give ranges of values according to which the patient population was stratified as shown in Table 6. Within each stratified range, the ratio of the number of subjects diagnosed with AD (McKhann et al., *Neurology* 34:939, 1984) to the number of subjects not having AD was determined, showing a correlation of increasing  $r$  value with increasing risk for AD (Table 6).

TABLE 6

Relative Risk for AD: by $r$ Value	
$r \times 100$	# of AD subjects: # of Non-AD subjects
>15	1.42
>20	1.8
>25	2.2
>30	2.0

## EXAMPLE 8

INCREASED PREDICTIVE VALUE OF EXMTDNA:MTDNA RATIO IN SUBJECTS HAVING  
AT LEAST ONE APOE4 ALLELE

Determination of the value  $r$  according to Example 7 was combined with 5 determination of APOE genotype according to Example 6 using the patient populations characterized in Examples 6 and 7 to ascertain the relationship of APOE genotype to relative risk for AD, and to demonstrate the increased correlative value of APOE genotype combined with increasing  $r$  values with relative risk for AD. The relative risk for AD as a function of APOE genotype using the APOE alleles 2, 3 or 4 alone is 10 shown in the bottom line of Table 7. The remainder of Table 7 shows the relative risk of AD as a function of APOE genotype and  $r$  value when subjects are stratified according to  $r$  as in Example 7. The relative risk for AD increases as a function of increasing  $r$ . In particular individuals with one APOE4 allele, and especially subjects who are homozygous for the APOE4 allele, exhibit increasing risk for AD as a function 15 of increasing  $r$ , as shown in Table 7.

TABLE 7

Relative Risk for AD: by $r$ value and ApoE Genotype (# AD: # Non-AD) ApoE Genotype					
$r \times 100$	e2/3	e3/3	one e4 allele	e3/4	e4/4
>15	1.5	1.5	2.2	2.1	7.1
>20	2.0	1.9	3.7	3.1	9.0
>25	1.7	1.8	5.3	3.0	>9.0
>30	2.5	2.0	4.6	3.0	>9.0
independent of $r$	0.5	0.6	1.9	1.7	6.3

## EXAMPLE 9

## IDENTIFICATION OF AGENTS THAT ALTER THE VALUE OF R

In this example, an agent suitable for treating AD is identified based on its ability to lower the value of  $r$  as defined above. A blood sample is obtained from a 5 patient diagnosed with AD (McKhann et al., *Neurology* 34:939, 1984) and a ratio  $r$  is calculated using the formula:

$$r = x/(x + y)$$

10 wherein

$x$  is the amount of exmtDNA in a sample, and

15  $y$  is the amount of mtDNA in the sample, where  $x$  and  $y$  are determined using the primer extension assay as described in Example 4. The candidate agent is then administered to the patient in a quantity and for a time sufficient to impart a therapeutically beneficial effect, and blood samples are periodically collected and processed using the primer extension assay as described in Example 4 to monitor alterations in the value  $r$ . Candidate agents are subjected to preliminary characterization for toxicity, bioavailability and modes of delivery prior to administration to a patient. An agent is selected that causes a reduction in the value for  $r$ , signifying an effect on the 20 altered mitochondrial function in the patient that may underlie the differential extractability of exmtDNA and mtDNA that contributes to  $r$  values associated with increased risk for AD.

## EXAMPLE 10

## 25 DETECTION OF EXMTDNA SEQUENCES IN TELOMERES

In this example, fluorescence *in situ* hybridization (FISH) is used to 30 localize nucleic acid sequences that are present in SEQ ID NO:1 to the telomeric region of at least one human chromosome. The methods of Cannizzaro et al. (*Methods Mol. Biol.* 75:313, 1997) and references cited therein, all of which are incorporated hereby in their entireties, are used to conduct FISH. Human cell lines are grown to confluence on

coverslips and then fixed and permeabilized for FISH analysis of metaphase chromosomes. The cloned pBeloBAC insert containing SEQ ID NO:1 as described above is excised from the plasmid vector and fluorescein labeled as described, and used to probe the fixed and permeabilized cells. Metaphase cells are evaluated by 5 fluorescent laser scanning confocal microscopy and subchromosomal localization of the fluorescent probe to telomeric regions is observed. Cytological methods are used to prepare metaphase spreads of human peripheral blood leukocytes as described in the cited references, and the chromosomes are doubly labeled using suitable reporter moieties for discerning two signals, one being quinacrine to identify each chromosome 10 by its characteristic banding pattern and the other being the labeled SEQ ID NO:1 insert, to correlate a particular chromosome with telomeric hybridization of the insert. The nucleic acid insert of SEQ ID NO:1 is next fragmented with restriction endonucleases, and the separated fragments are labeled to generate a panel of probes representing distinct sequence portions of SEQ ID NO:1, to determine whether the 15 portion of the insert hybridizing to telomeres corresponds to an exmtDNA sequence or a non-exmtDNA sequence present in the insert as it was cloned from the human genomic library, as described above in Example 1.

#### EXAMPLE 11

##### 20 DETECTION OF EXMTDNA SEQUENCES IN TELOMERES BY RADIATION HYBRID MAPPING

This example describes radiation hybrid (RH) screening by PCR using the Stanford G3 panel (Research Genetics, Huntsville, AL) to positionally map exmtDNA sequences in the human genome. Localization of the exmtDNA sequences within the nuclear DNA was accomplished by radiation hybrid (RH) mapping according to 25 Boehnke *et al.* (1991 *Am. J. Hum. Genet.* 49:1174) and Walter *et al.* (1994 *Nat. Genet.* 7:22). The Stanford G3 RH panel consists of more than eighty whole-genome radiation hamster-human hybrids and 14,000 characterized chromosomal markers available at the Stanford Human Genome Center (SHGC; <http://www-shgc.stanford.edu>). Eighty-three hybrid clones of the G3 RH panel, plus both parental human and hamster control cells,

were screened with PCR primers homologous to a portion of the CO III gene region within the exmtDNA sequence. PCR primers were designed to select for the exmtDNA sequence and contained nucleotide changes at positions corresponding to nucleotides 9325, 9329, 9335, 9540, 9545, and 9548 of SEQ ID NO:2:

5

L-strand primer: 5'- TCCACTCCACAAACCCTCCTT SEQ ID NO:109

H-strand primer: 5'- GGGCCAGTGTCCCCTAG SEQ ID NO:110

10 Human and hamster genomic DNAs (Research Genetics, Huntsville, AL) were  
used to confirm specific amplification of a 223 bp fragment (exmtDNA marker) using  
AmpliTaq® DNA polymerase (Perkin-Elmer, Norwalk, CT) and standard PCR  
conditions as described above. Analysis was performed on the RH Server operated by  
the Stanford Human Genome Center (SHGC; <http://www-shgc.stanford.edu>, Palo Alto,  
15 CA). The exmtDNA marker was linked to more than 5,000 framework markers and  
EST's that have been characterized by the SHGC and by Genethon  
([http://www.genethon.fr/genethon\\_en.html](http://www.genethon.fr/genethon_en.html)).

An amplification product of the expected 223 bp size was obtained from hybrid clone numbers 3, 11, 33, 40, 49, 50, 59, 64, 73, and 75, as well as from the human control cells, and not from the hamster control cells. The closest SHGC markers identified were SHGC-57364 with positive hybrid clone numbers 3, 11, 33, 40, 49, 50, 59, 64, 73, and 75; SHGC-4723 with positive hybrid clone numbers 11, 33, 49, 50, 51, 59, 64, 68, 73, and 75; and SHGC-428 with positive hybrid clone numbers 9, 11, 40, 49, 50, 51, 59, 64, 68, 73, and 75. All three markers were localized to a very distal portion at or near the telomeric region of chromosome 1, at distances from the exmtDNA of 0, 26, and 32 centirays (cR; one cR corresponds roughly to 26 kb), which is roughly equivalent to 0, 676, and 832 kb, respectively (Figure 5). SHGC-4723 and SHGC-428 are ordered markers linked to Bin#1 of chromosome 1 at GDB Locus D1S2565 and D1S243, respectively. SHGC-57364 is a positional marker, also linked to Bin#1 of chromosome 1, that has been oriented at a distance of 26 cR from SHGC-4723 (LOD

7.11). Interestingly, this marker (GenBank Accession No. T10998) originated from a 230 nucleotide sequence with partial similarity to the mitochondrial gene for cytochrome *c* oxidase subunit 2. It is therefore conceivable that the PCR primers used for identifying this marker in the clones where it was found may also have amplified the 5 exmtDNA sequence. However, independent confirmation of the exmtDNA co-localization was provided by the specific CO III primer set described herein. These data demonstrate unambiguously the nuclear chromosomal location of this exmtDNA and its presence as a single copy sequence.

10

#### EXAMPLE 12

##### DETECTION OF EXMTDNA BY SOUTHERN BLOT HYBRIDIZATION ANALYSIS

The nucleotide substitution A→C at position 6266 in the portion of the exmtDNA sequence corresponding to the CO I gene (Fig. 4) creates a six-base pair *Nae* I restriction recognition site which is not present in human wild type mtDNA. In this 15 example, this difference is exploited to distinguish mtDNA from the exmtDNA sequence by Southern blot analysis of human cellular DNA isolated from white blood cells (Figure 6a) or from ρ<sup>0</sup> cells (Fig. 6b), or of mtDNA from isolated human brain mitochondria (Fig. 6b).

DNA from two normal control individuals was restricted with *Sac* I, which 20 cleaved at nucleotide position 9647 in the CO II gene of both the mtDNA and exmtDNA, and with *Nae* I, which cleaved at nucleotide position 932 in the 12S rRNA gene of mtDNA and nucleotide position 6266 in the exmtDNA. A wild type mitochondrial ATPase 8 probe detected the resulting 8715 bp mtDNA and 3381 bp exmtDNA fragments. The signal intensities of the two DNA fragments indicated their 25 relative abundance in the cellular DNA preparation, suggesting a low copy number or possibly a single copy of this exmtDNA within the genomic DNA.

Human brain samples were obtained post mortem as soon as possible after death. The biological samples were frozen and stored at -80°C. and shipped on dry ice to analysis facilities. Mitochondria were purified from human brain tissue and mtDNA

was extracted as described by Mecocci *et al.* (1994 *Ann. Neurol.* 36:747-751). Briefly, human parietal cortex tissue was homogenized in an isotonic solution (0.05 M Tris-HCl, pH 8.0, 0.21 M mannitol, 0.07 M sucrose, 3 mM calcium chloride) using a glass dounce homogenizer and a round bottom glass rod with 0.15 µm clearance. EDTA concentration was brought to 0.02 M and the homogenate was centrifuged at 1,500 g for 20 min at 4°C. The supernatant was transferred to a new tube, centrifuged again at 1,500 g for 20 min at 4°C, then transferred to a new tube and centrifuged at 17,000 g for 20 min at 4°C. The pellet consisting of mitochondria was resuspended in 40 mM Tris-HCl, pH 8.0, 10 mM magnesium chloride, containing DNase I and RNase A (Roche Molecular Biochemicals, Inc., Indianapolis, IN) at concentrations of 40 units/ml and 0.02 mg/ml, respectively, and incubated for 45 min at 37°C. The organelle suspension was diluted to 20 ml with isotonic dilution buffer (0.05 M Tris-HCl, pH 8.0, 0.21 M mannitol, 0.07 M sucrose, 0.01 M EDTA) and centrifuged at 17,000 g for 20 min at 20°C. The pellet was resuspended in isotonic dilution buffer, combined with an equal volume of lysis buffer (0.01 M Tris-HCl, pH 8.0, 0.1 M NaCl, 1 mM EDTA, 0.5% SDS) containing 400 µg/ml Proteinase K (Roche Molecular Biochemicals, Inc., Indianapolis, IN) and incubated for 3 hr at 37°C. DNA was purified by three phenol extractions followed by three chloroform:isoamyl alcohol (24:1) extractions and ethanol precipitation. The DNA was resuspended in buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA) and its concentration was determined by UV absorbance at 260 nm.

Total DNA isolated from white blood cells with DNAzol reagent, DNA isolated from  $\rho^0$  118/5 SH-SY5Y cells with DNAzol reagent, and DNA extracted from purified human brain mitochondria as described above were treated with *Nae* I and *Sac* I restriction endonucleases (New England Biolabs, Beverly, MA), separated by horizontal agarose gel electrophoresis, and blotted onto Nytran Plus membranes (Schleicher & Schuell, Keene, NH). Membranes were probed with a ~200 bp fragment of the wildtype ATP synthase subunit 8 (ATPase 8) gene that had been generated by PCR amplification of a DNA preparation from white blood cells using the following oligonucleotide primers:

L-strand primer: 5'-TAGCATT AACCTTTAAGTTAAAGA      SEQ ID NO:111

H-strand primer: 5'-TCGTTCA TTTGGTTCTCA      SEQ ID NO:112

5        The probe was labeled with  $^{32}\text{P}$ -dATP (Amersham, Arlington Heights, IL) and the random priming method (Prime-a-Gene<sup>TM</sup> Labeling System, Promega, Madison, WI) according to the suppliers' recommendations. The blot was incubated with approximately 50 million cpm (5-10 ng) of probe for 16 h at 42°C in hybridization solution (6X SSC, 50% formamide, 1X Denhardt's solution, 1% SDS and 100  $\mu\text{g}/\text{ml}$  salmon sperm DNA). Blots were washed twice for 5 min in 6X SSC, 0.1% SDS at room temperature, once for 30 min in 2X SSC, 0.1% SDS at 42°C, and once for 30 min in 0.2X SSC, 0.1% SDS at 42°C, and exposed to XAR 5 film (Kodak) for 1 to 4 days and developed.

Figure 6a shows hybridization of the ATPase 8 probe to total DNA isolated with DNazol reagent from white blood cells of two control individuals (A, B), and Figure 6b shows probe hybridization to DNA isolated from  $\rho^0$  SH-SY5Y cells with DNazol reagent and from purified human brain mitochondria with Proteinase K digestion, as described above. Approximately 10  $\mu\text{g}$  of each DNA sample was restricted with *Nae* I and *Sac* I, and probed with the PCR amplified ATPase 8 probe.

20 Wildtype mtDNA carries a *Nae* I site at nucleotide position 932 which yields a 8,715 bp restriction fragment in combination with *Sac* I restriction at nucleotide position 9647. *Nae* I recognized a restriction site present in the exmtDNA sequence, which was created by the nucleotide substitution at position 6266 (Fig. 4) and was not present in the wildtype mtDNA. By cleavage of the exmtDNA at this site and at nucleotide position 25 9647 by *Sac* I, a 3,381 bp restriction fragment was created. An unlabeled aliquot (~4 ng) of the 200 bp ATPase 8 probe (Fig. 6a, probe control) showed an appropriate hybridization signal. Exposure times of the blots to X-ray film were (Fig. 6a) 24 hr and (Fig. 6b) 4 d.

To provide additional support for an extramitochondrial (e.g., nuclear) origin of the 5.8 kb exmtDNA sequence of SEQ ID NO:1, DNA from human  $\rho^0$  SH-

SY5Y cells (*i.e.*,  $\rho^0$  cells depleted of mtDNA as described in Miller et al., 1996 *J. Neurochem.* 67:1897) was restricted with *Sac* I and *Nae* I and hybridized with the wildtype mitochondrial ATPase 8 probe, which detected the resulting 3381 bp exmtDNA fragment (Figure 6b). To further demonstrate that the exmtDNA sequence 5 does not reside within mitochondria, DNA was extracted from purified mitochondria derived from a human parietal cortex brain sample and digested with *Sac* I and *Nae* I, as described above. Following Southern blot hybridization with the wildtype mitochondrial ATPase 8 probe, an 8715 bp mtDNA fragment was detected but no detectable 3381 bp exmtDNA fragment was present (Figure 6b).

10

#### EXAMPLE 13

##### DETECTION OF TELOMERIC CHANGES USING EXMTDNA SEQUENCES

The nucleotide sequences of the 5.8 exmtDNA (SEQ ID NO:1), and its flanking regions, are used to design nucleic acid probes that specifically detect nucleic acids 15 corresponding to the telomere region located at one end of human chromosome 1. Methods utilizing such probes are used to determine the degree of telomeric shortening of chromosome 1.

A synthetic oligonucleotide comprising a junction sequence of the exmtDNA is labeled and used as a probe of nuclear DNA in Southern blot analysis under stringent 20 conditions, or as an unlabeled primer for PCR amplification of the telomeric region of chromosome 1. In PCR amplifications, the second primer is labeled and has a sequence specific for the telomere repeat sequence (5'-TTAGGG in humans); such primers are commercially available from Roche Molecular Biochemicals (Indianapolis, IN). The labeled PCR products are detected and their length assayed according to methods 25 known in the art (*e.g.*, polyacrylamide gel electrophoresis (PAGE) with appropriate molecular weight markers). While not wishing to be bound by any theory, somatic cells from more aged individuals may have fewer numbers of telomeric repeats on the ends of their chromosomes than those from less aged individuals and, accordingly, the lengths of the PCR amplification products in biological samples derived from these

more aged individuals are shorter than those in biological samples derived from younger individuals. In like fashion, because some tumor cells may reactivate the enzyme (telomerase) that promotes the addition of telomeric repeats in germline (but not, typically, in somatic) cells, the lengths of the PCR amplification products in 5 biological samples derived from a tumor are longer than those in biological samples derived from healthy individuals.

From the foregoing, it will be appreciated that, although specific 10 embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

## CLAIMS

1. A method for determining the risk for or presence of a disease associated with altered mitochondrial function in a first subject suspected of having or being at risk for having such a disease, comprising:

comparing a ratio  $r$  for each of a first and a second biological sample containing extramitochondrial DNA and mitochondrial DNA, said first biological sample being obtained from said first subject and said second sample being obtained from a second subject known to be free of a risk or presence of a disease associated with altered mitochondrial function, using the formula:

$$r = x/(x + y)$$

wherein

$x$  is the amount of extramitochondrial DNA in a sample, and

$y$  is the amount of mitochondrial DNA in the sample; and

therefrom determining the risk or presence of the disease.

2. The method of claim 1 wherein the ratio  $r$  is calculated by a method comprising:

contacting a biological sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said extramitochondrial DNA and present in said mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA.

3. The method of claim 1 wherein the ratio  $r$  is calculated by a method comprising:

contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said amplified extramitochondrial DNA and present in said amplified mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA.

4. The method of claim 1 wherein the ratio  $r$  is calculated by a method comprising:

contacting a biological sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said extramitochondrial DNA and present in said mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from said first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA.

5. The method of claim 1 wherein the ratio  $r$  is calculated by a method comprising:

contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said amplified extramitochondrial DNA and present in said amplified mitochondrial DNA, under conditions and for a time sufficient to allow

hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from said first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA.

6. The method of claim 1 wherein the biological sample is treated by heating said sample in water to lyse cells contained in the sample, and then extracting cellular DNA from said lysed cells using an aqueous DNA extraction procedure.

7. The method of claim 6 wherein the sample comprises a crude buffy coat fraction of whole blood.

8. The method of claim 1, further comprising the step of determining the ApoE genotype of the first subject and correlating said genotype with the risk or presence of disease.

9. The method of claim 1 wherein the disease associated with altered mitochondrial function is selected from the group consisting of Alzheimer's Disease, Huntington's Disease, Parkinson's Disease, dystonia, schizophrenia, non-insulin dependent diabetes mellitus, mitochondrial encephalopathy, lactic acidosis, and stroke, myoclonic epilepsy ragged red fiber syndrome, and Leber's hereditary optic neuropathy.

10. A method for quantifying extramitochondrial DNA, comprising:  
contacting a sample containing extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence complementary to at least a portion of said extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA; and

detecting hybridization of the primer to the extramitochondrial DNA, and therefrom quantifying the extramitochondrial DNA.

11. A method for quantifying extramitochondrial DNA, comprising:  
contacting a sample containing extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence complementary to at least a portion of said extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA; and

detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a product, and therefrom quantifying the extramitochondrial DNA.

12. A method for quantifying extramitochondrial DNA, comprising:  
contacting a sample containing amplified extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence complementary to at least a portion of said extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA; and

detecting hybridization of the primer to the extramitochondrial DNA, therefrom quantifying the extramitochondrial DNA.

13. A method for quantifying extramitochondrial DNA, comprising:  
contacting a sample containing amplified extramitochondrial DNA with an oligonucleotide primer having a nucleotide sequence complementary to at least a portion of said extramitochondrial DNA under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA; and  
detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a product, and therefrom quantifying the extramitochondrial DNA.

14. The method of claim 12 wherein the extramitochondrial DNA is amplified using a technique selected from the group consisting of polymerase chain reaction, transcriptional amplification systems and self-sustained sequence replication.

15. The method of any one of claims 1-14 wherein a single oligonucleotide primer is used.

16. The method of any one of claims 1-14 wherein a primer extension assay is used.

17. The method of either claim 12 or 14 wherein the step of detecting comprises a technique selected from the group consisting of polymerase chain reaction, primer extension assay, ligase chain reaction, and restriction fragment length polymorphism analysis.

18. The method of any one of claims 1, 10, 11, 12 or 13 wherein the amount of extramitochondrial DNA in the biological sample is quantified by determining the presence in said sample of a nucleotide sequence selected from the group consisting of:

- (a) SEQ ID NO:1,
- (b) a portion of SEQ ID NO:1,
- (c) SEQ ID NO:3,
- (d) a portion of SEQ ID NO:3,

(e) an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one nucleotide substitution of Figure 4 at a corresponding nucleotide position, and

(f) an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one nucleotide deletion of Figure 4 at a corresponding nucleotide position.

19. The method of claim 18 wherein the portion of the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial cytochrome c oxidase.

20. The method of claim 18 wherein the portion of SEQ ID NO:1 corresponds to a portion of a mitochondrial cytochrome c oxidase encoding sequence selected from the group consisting of a portion of a CO1 encoding sequence and a portion of a CO2 encoding sequence.

21. The method of claim 18 wherein the portion of the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial ATP synthetase subunit.

22. The method of claim 18 wherein the portion of SEQ ID NO:1 corresponds to a portion of a mitochondrial ATP synthetase subunit encoding sequence selected from the group consisting of a portion of a sequence encoding ATP synthetase subunit 6 and a portion of a sequence encoding ATP synthetase subunit 8.

23. The method of claim 18 wherein the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of SEQ ID NO:2 selected from the group consisting of a portion of a sequence encoding ND1, a sequence encoding a portion of ND2 and a sequence encoding a portion of CO3.

24. The method of claim 18 wherein the portion of the nucleotide sequence of SEQ ID NO:3 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial ATP synthetase subunit.

25. The method of claim 18 wherein the portion of SEQ ID NO:3 corresponds to a portion of a mitochondrial ATP synthetase subunit encoding sequence selected from the group consisting of a portion of a sequence encoding ATP synthetase subunit 6 and a portion of a sequence encoding ATP synthetase subunit 8.

26. The method of claim 18 wherein the portion of the nucleotide sequence of SEQ ID NO:1 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial tRNA.

27. The method of claim 18 wherein the portion of the nucleotide sequence of SEQ ID NO:3 corresponds to a portion of the nucleotide sequence of SEQ ID NO:2 encoding a mitochondrial tRNA.

28. An isolated nucleic acid comprising all or a portion of the nucleotide sequence of SEQ ID NO:1 or a complementary sequence thereto.

29. An isolated nucleic acid comprising all or a portion of a nucleotide sequence of SEQ ID NO:1 or a complementary sequence thereto, wherein the sequence of said isolated nucleic acid differs by at least one nucleotide from the corresponding sequence of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:2 or a complementary sequence thereto.

30. An isolated nucleic acid comprising all or a portion of the nucleotide sequence of SEQ ID NO:3 or a complementary sequence thereto.

31. An isolated nucleic acid comprising all or a portion of a nucleotide sequence of SEQ ID NO:3 or a complementary sequence thereto, wherein the sequence of said isolated nucleic acid differs by at least one nucleotide from the corresponding sequence of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:2 or a complementary sequence thereto.

32. A method for determining the risk or presence of a disease associated with altered mitochondrial function in a subject suspected of having or being at risk for having such a disease, comprising:

quantifying the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from said subject, and therefrom determining the risk or presence of the disease.

33. A method for determining the risk or presence of a disease associated with altered mitochondrial function in a first subject suspected of having or being at risk for having such a disease, comprising:

comparing the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from said first subject to the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from a second subject, and therefrom determining the risk or presence of the disease.

34. A method for determining the risk or presence of a disease associated with altered mitochondrial function in a first subject suspected of having or being at risk for having such a disease, comprising:

quantifying the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from said subject and comparing said amount of extramitochondrial DNA and said amount of mitochondrial DNA to the amount of extramitochondrial DNA and the amount of mitochondrial DNA in a biological sample from a second subject known to be free of a risk or presence of a disease associated with altered mitochondrial function, and therefrom determining the risk or presence of the disease.

35. A method of monitoring a telomere comprising:

detecting in a telomeric region the presence of a nucleic acid molecule comprising all or a portion of SEQ ID NO:1 or a complementary portion thereto.

36. The method of claim 35 wherein the detected nucleic acid molecule comprises an exmtDNA sequence.

37. A method of monitoring a telomere comprising:  
detecting in a telomeric region the presence of a nucleic acid molecule comprising all or a portion of SEQ ID NO:3 or a complementary portion thereto.

38. A method of identifying an agent suitable for treating a disease associated with altered mitochondrial function, comprising:

comparing a ratio  $r$  from a sample obtained before contacting a biological source with a candidate agent to the ratio  $r$  from a sample obtained after contacting the biological source with the candidate agent, said ratio  $r$  calculated using the formula:

$$r = x/(x + y)$$

wherein

$x$  is the amount of extramitochondrial DNA in a sample, and

$y$  is the amount of mitochondrial DNA in the sample; and

therefrom determining the suitability of said candidate agent for treating a disease associated with altered mitochondrial function.

39. The method of claim 38 wherein the biological sample comprises a crude buffy coat fraction of whole blood.

40. The method of claim 39 wherein the biological sample is treated by heating said sample in water to lyse cells contained in the sample, and then extracting cellular DNA from said lysed cells using an aqueous DNA extraction procedure.

41. The method of claim 38 wherein the ratio  $r$  is calculated by a method comprising:

contacting a sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said extramitochondrial DNA and present in said mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio  $r$ .

42. The method of claim 38 wherein the ratio  $r$  is calculated by a method comprising:

contacting a sample containing extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said extramitochondrial DNA and present in said mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from said first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio.

43. The method of claim 38 wherein the ratio  $r$  is calculated by a method comprising:

contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said amplified extramitochondrial DNA and present in said amplified mitochondrial DNA, under conditions and for a time sufficient to allow

(e) an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one nucleotide substitution of Figure 4 at a corresponding nucleotide position, and

(f) an extramitochondrial DNA sequence comprising a nucleic acid sequence that (i) corresponds to at least a portion of SEQ ID NO:2 and (ii) contains at least one nucleotide deletion of Figure 4 at a corresponding nucleotide position.

46. The method of claim 45 wherein the nucleotide sequence of SEQ ID NO:1 or a portion thereof corresponds to a mitochondrial cytochrome c oxidase encoding sequence of SEQ ID NO:2 or a portion thereof.

47. The method of claim 49 wherein the mitochondrial cytochrome c oxidase encoding sequence of SEQ ID NO:2 or a portion thereof is selected from the group consisting of a sequence encoding CO I or a portion thereof and a sequence encoding CO2 or a portion thereof.

48. The method of claim 45 wherein the nucleotide sequence of SEQ ID NO:1 or portion thereof, or the nucleotide sequence of SEQ ID NO:3 or portion thereof corresponds to a mitochondrial ATP synthetase subunit encoding sequence of SEQ ID NO:2 or a portion thereof.

49. The method of claim 48 wherein the mitochondrial ATP synthetase subunit encoding sequence of SEQ ID NO:2 or a portion thereof is selected from the group consisting of a sequence encoding ATP synthetase subunit 6 or a portion thereof and a sequence encoding ATP synthetase subunit 8 or a portion thereof.

50. The method of claim 45 wherein the nucleotide sequence of SEQ ID NO:1 corresponds to a sequence of SEQ ID NO:2 or a portion thereof selected from the group consisting of a sequence encoding a truncated NADH dehydrogenase subunit 1 or a portion

hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization of the primer to the extramitochondrial DNA and to the mitochondrial DNA, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio  $r$ .

44. The method of claim 38 wherein the ratio  $r$  is calculated by a method comprising:

contacting a sample containing amplified extramitochondrial DNA and mitochondrial DNA with an oligonucleotide primer having a nucleotide sequence that is complementary to a sequence present in said amplified extramitochondrial DNA and present in said amplified mitochondrial DNA, under conditions and for a time sufficient to allow hybridization of said primer to the extramitochondrial DNA and to the mitochondrial DNA; and

detecting hybridization and extension of the primer to the extramitochondrial DNA to produce a first product and hybridization and extension of the primer to the mitochondrial DNA to produce a second product distinguishable from said first product, and therefrom quantifying the extramitochondrial DNA and the mitochondrial DNA to calculate the ratio.

45. The method of claim 38 wherein comparing the ratio  $r$  from a sample obtained before contacting a biological source with a candidate agent to the ratio  $r$  from a sample obtained after contacting the biological source with the candidate agent comprises determination of the presence in said sample of a nucleotide sequence selected from the group consisting of:

- (a) SEQ ID NO:1,
- (b) a portion of SEQ ID NO:1,
- (c) SEQ ID NO:3,
- (d) a portion of SEQ ID NO:3.

thereof, a sequence encoding NADH dehydrogenase subunit 2 or a portion thereof and a sequence encoding truncated CO3 or a portion thereof.

51. The method of claim 38 wherein the disease associated with altered mitochondrial function is selected from the group consisting of Alzheimer's Disease, Huntington's Disease, Parkinson's Disease, dystonia, schizophrenia, non-insulin dependent diabetes mellitus, mitochondrial encephalopathy, lactic acidosis, and stroke, myoclonic epilepsy ragged red fiber syndrome, and Leber's hereditary optic neuropathy.

52. A method of identifying an agent suitable for treating a subject suspected of being at risk for having a disease associated with altered mitochondrial function, comprising:

determining the apolipoprotein E genotype of the subject;

comparing a ratio  $r$  in a biological sample obtained from the subject before contacting said sample with a candidate agent to the ratio  $r$  in a biological sample obtained from the subject after contacting said sample with a candidate agent, said ratio  $r$  calculated using the formula:

$$r = x / (x + y)$$

wherein

$x$  is the amount of extramitochondrial DNA in the sample, and

$y$  is the amount of mitochondrial DNA in the sample; and

therefrom determining the suitability of said candidate agent for treating the disease associated with altered mitochondrial function.

53. The method of claim 52 wherein the disease associated with altered mitochondrial function is Alzheimer's disease.

54. A method of correlating a ratio  $r$  with the suitability of an agent for treating Alzheimer's disease in a subject, comprising:

determining a ratio  $r$  in a biological sample obtained from the subject, said ratio  $r$  calculated using the formula:

$$r = x / (x + y)$$

wherein

$x$  is the amount of extramitochondrial DNA in the sample, and

$y$  is the amount of mitochondrial DNA in the sample;

contacting said subject with a candidate agent and evaluating the subject for alterations in the AD disease state, and therefrom correlating the suitability of the agent for treating AD in the subject with  $r$ .

55. The method of claim 54, further comprising determining the apolipoprotein E genotype of the subject, and therefrom correlating the suitability of the agent for treating AD in the subject with  $r$  and with the apolipoprotein E genotype.

cactgtaaag ctaacccagc attaaccttt taagttaaag actaagagaaa tcattatctc	60
tttacagtga aatgccacag ctaaaatacca ctgttatgacc tgcttatcatc accccaataac	120
tcctcacgtt atttctcatc acccaactaa aaatactaaa cacacactgc catctgccc	180
cctcaccaaa atttattaaa ataaaaaaact acagtaagcc ctgagaacca aaatgaacga	240
aaatttattc gtttcattca ttacccctac agtacttaggc ctacccgcca cagtaccaat	300
catccttattt ccccccttac tggtcccaac ctccaaatac ctcataaca accgactaat	360
caccactcaa caatgactac ttcaactcac cttaaaacaa ataataacga tacataacat	420
taaggacga acctggtccc ttatactaattt ttccctgatt atttttattt ccacaactaa	480
tctcctcgga ctcttgcccc actcatttac accaattcac	519

*Fig. 1*

*Fig. 2A*

gatcacaggt	ctatcaccct	attaaccact	cacgggagct	ctccatgcac	ttggtat	tttt	60
cgtctgggg	gtatgcacgc	gatagcattt	cgagacgctg	gagccggagc	accctatgtc		120
gcagtatctg	tctttgattt	ctgcctcattc	ctattatttt	tcgcacctac	gttcaatatt		180
acaggcgaac	atacttacta	aagtgtgtt	attaatttt	gcttgttagga	cataataata		240
acaattgaat	gtctgcacag	ccacttcca	cacagacatc	ataacaaaaaa	atttccacca		300
aaccccccct	cccccgctt	tggccacagc	acttaaacac	atctctgcca	aaccccaaaa		360
acaaaagaacc	ctaaccaccag	cctaaccaga	tttcaaattt	tatcttttg	cggtatgcac		420
tttaaacagt	caccccccac	ctaaccacatt	atttccctt	cccaactccca	tactactaat		480
ctcatcaata	caaccccccgc	ccatcctacc	cagcacacac	acaccgctgc	taacccata		540
ccccgaacca	accaaaccccc	aaagacaccc	cccacagttt	atgtagctt	cctcctcaaa		600
gcaataacact	gaaaatgtt	agacgggctc	acatcacccc	ataaaacaaat	aggtttggc		660
ctagccttc	tattagctct	tagtaagatt	acacatgca	gcatccccgt	tccagttagt		720
tcaccctcta	aatcaccacg	atcaaaagg	acaagcatca	agcacgcagc	aatgcagctc		780
aaaacgccta	gcctagccac	accccccacgg	gaaacagcag	tgattaacct	ttagcaataa		840
acgaaagttt	aactaagctt	tactaacc	agggttggc	aatttctgtc	cagccaccgc		900
ggtcacacga	ttaacccaag	tcaatagaag	ccggcgtaaa	gagtgtttt	gatcaccccc		960
tcccaataaa	agctaaaact	cacctgagtt	gtaaaaaaact	ccagttgaca	aaaaatagac		1020
tacgaaagtg	gctttaacat	atctgaacac	acaatagctt	agacccaaac	tgggattaga		1080
taccccaacta	tgttagccc	taaacctcaa	cagttaaatc	aacaaaactg	ctcgccagaa		1140
cactacgagc	cacagctaa	aactcaaagg	acctggcggt	gcttcataatc	cctctagagg		1200
agcctgtct	gtaatcgata	aaccccgatc	aacctccatc	cctcttgctc	agcctatata		1260
ccgcatctt	cagcaaacc	tgatgaaggc	tacaaagtaa	gcbcgaagtac	ccacgtaaag		1320
acgtaggtc	aagggttagc	ccatgaggtg	gcaagaaaatg	ggctacattt	tctaccccaq		1380
aaaactacga	tagcccttat	gaaacttaag	ggtcgaaggt	ggatttagca	gtaaactaaag		1440
agtagagtgc	ttagttgaac	agggccctga	agcgcgtaca	caccgcgcgt	caccctcctc		1500
aagtataact	caaaggacat	ttaactaaaa	cccttacgc	tttatataga	ggagacaagt		1560
cgtAACATGG	taagtgtact	ggaaagtgc	cttggacgaa	ccagagtgt	gcttaacaca		1620
aagcacccaa	tttacactta	ggagatttca	acttaacttgc	accgctctga	gctaaaccta		1680
gccccaaacc	cactccaccc	tactaccaga	caaccttagc	caaaccattt	acccaaataa		1740
agtataggcg	atagaaattt	aaacctggcg	caatagatat	agtaccgca	gggaaagatg		1800
aaaaattata	accaaggcata	atatacgaa	gactaacc	tataccttct	gcataatgaa		1860
ttaactagaa	ataactttgc	aaggagagcc	aaagctaaga	cccccgaaac	cagacgagct		1920
acctaagaac	agctaaaaga	gcacacccgt	ctatgttagc	aaatagtgg	aagattata		1980
ggtagaggcg	acaaaacctac	cgagcctggt	gatagctgtt	tgtccaagat	agaatcttag		2040
ttcaacttta	aatttgc	cagaacc	taaattccct	tgtaaattt	actgttagt		2100
caaagaggaa	cagcttttgc	gacacttagg	aaaaaccttgc	tagagagat	aaaaaatttta		2160
acacccatag	taggcctaaa	agcagccacc	aattaagaaa	gcgttcaagc	tcaacaccc		2220
ctacctaaaa	aatcccaaaac	atataactga	actcctcaca	cccaatttgg	ccaatctatc		2280
accctataga	agaactaatgt	tttagtataag	taacatgaaa	acattctcc	ccgcataagc		2340
ctgcgtcaga	ttaaaaacact	gaactgacaa	ttaacagccc	aatatctaca	atcaaccaac		2400
aagtattat	taccctcact	gtcaacccaa	cacaggcatg	ctcataagga	aaggtttaaa		2460
aaagtaaaag	gaactcggc	aatcttaccc	cgcctgttgc	ccaaaaacat	cacctctagc		2520
atcaccatgt	tttagaggc	cgccctggcc	gtgacacatg	ttaacggcc	gcgggtaccc		2580
aaccgtgca	aggttagcata	atcacttgc	ccttaatag	ggacctgtat	gaatggctcc		2640
acgagggttc	agctgtctc	tacttttac	cagtgttttgc	gacctgccc	tgaagaggcg		2700
ggcataaacac	agcaagacga	gaagaccctt	tggagcttgc	atttattaaat	gcaaacagta		2760
cctaacaac	ccacagggtcc	taaactacca	aacctgcatt	aaaaatttgc	gttggggcga		2820
cctcgagca	gaacccaaacc	tccgagcgt	acatgctaag	acttcaccag	tcaaagcgaa		2880
ctactatact	caattgtatcc	aataacttgc	ccaacggaaac	aagttaccct	agggataaca		2940
gcfgcaatctt	attcttagat	ccatatcaac	aatagggttt	acgacctcg	tgttggatca		3000
ggacatcccg	atgggtc	cgctatttt	gttgcgttt	ttaacacgtt	aaagtcctac		3060
gtgatctgag	ttcagaccgg	agtaatccag	gtcgggttct	atctaccc	aaattccctcc		3120
ctgtacgaaa	ggacaagaga	aataaggcc	acttcacaaa	gccccttccc	ccgtaaaatga		3180
tatcatctca	acttagtatt	ataccacac	ccacccaaga	acagggttttgc	ttaagatggc		3240
agagcccggt	aatcgat	aactaaaac	tttacagtca	gaggttcaat	tcctcttctt		3300
aacaacatac	ccatggccaa	cctctactc	ctcattgtac	ccattctaaat	cgcaatggca		3360
ttcctaattgc	ttaccgaaac	aaaaatttca	ggctatatac	aactacgca	aggccccaaac		3420
tttgcgttgc	cctacgggt	actacaaccc	ttcgcgtac	ccataaaaact	tttcacccaa		3480

*Fig. 2B*

3/42

gagcccctaa	aacccgcccc	atctaccatc	accctctaca	tcaccggccc	gaccttagct	3540
ctcaccatcg	ctcttctact	atgaacccccc	ctccccatac	ccaacccct	ggtcaacctc	3600
aacctaggcc	tccttatttat	tctagccacc	tctagcttag	cggtttactc	aatctctga	3660
tcagggtgag	catcaaactc	aaactacgccc	ctgatcgccg	cactgcgagc	agttagccaa	3720
acaatctcat	atgaagtac	cctagccatc	attctactat	caacattact	aataagtggc	3780
tccttaacc	tctccacccct	tatcacaaca	caagaacacc	tctgattact	cctgccatca	3840
tgacccttgg	ccataatatg	atttatctcc	acactagcag	agaccaacccg	aaccccttc	3900
gaccttgcgg	aaggggagtc	cgaactagtc	tcaggcttca	acatcgaaa	cgccgcaggc	3960
ccttcgccc	tatttctcat	agccgaatac	acaaacatta	ttataataaa	caccctcacc	4020
actacaatct	tccttaggaac	aacatatgac	gcacttccc	ctgaactcta	cacaacatata	4080
tttgtcacca	agaccctact	tctaacctcc	ctgttcttat	gaattcgaac	agcatacccc	4140
cgattccgct	acgaccaact	catacaccc	ctatgaaaaa	acttcctacc	actcacccta	4200
gcattactta	tatgatatgt	ctccatacc	attacaatct	ccagcattcc	ccctcaaacc	4260
taagaaatat	gtctgataaa	agagttactt	tgatagagta	aataatagaa	gcttaaacc	4320
ccttattttct	aggactatga	gaatcgaaacc	catccctgag	aatccaaaat	tctccgtgcc	4380
acctatcaca	ccccatccta	aagtaaggc	agctaaaataa	gctatcgcc	ccataccccc	4440
aaaatgttgg	ttataccctt	cccgtaactaa	ttaatcccct	ggcccaaccc	gtcatctact	4500
ctaccatctt	tgcaggcaca	ctcatcacag	cgctaagctc	gcactgattt	tttacctgag	4560
taggcttaga	aataaaacatg	ctagctttt	ttccaggttct	aaccaaaaaaa	ataaaaccctc	4620
gttccacaga	agctgccatc	aagtatttcc	tcacgcagc	aaccgcatcc	ataatccttc	4680
taatagctat	cctcttcaac	aatatactct	ccggacaatg	aaccataacc	aataactacca	4740
atcaatactc	atcattaaata	atcataatag	ctatagcaat	aaaactagga	atagccccct	4800
ttcacttctg	agtcccagag	gttacccaaag	gcaccctct	gacatccg	ctgcttcttc	4860
tcacatgaca	aaaactagcc	cccatctcaa	tcatatacc	aatctctccc	tcactaaacg	4920
taagccttct	cctcactctc	tcaatcttat	ccatcatagc	aggcagttga	ggtggattaa	4980
accagaccca	gctacgcaaa	atcttagcat	actcctcaat	tacccacata	ggatgaataa	5040
tagcaattct	accgtacaac	ccttaacataa	ccattcttaa	ttaactatt	tatattatcc	5100
taactactac	cgcattctta	ctactcaact	taaactccag	caccacgacc	ctactactat	5160
ctcgcacctg	aaacaagacta	acatgactaa	cacccttaat	tccatccacc	ctccctccc	5220
taggaggcct	gcccccgcta	accggctttt	tgcccaaatg	ggccattatc	gaagaattca	5280
caaaaaacaa	tagcctcatc	atccccacca	tcatagccac	catcaccctc	cttaacctct	5340
acttctacct	acgcctaatac	tactccacct	caatcacact	actccccata	tctaacaacg	5400
taaaaaataaa	atgacagttt	gaacatacaa	aacccacccc	attcctccccc	acactcatcg	5460
cccttaccac	gctactctta	cctatctccc	ctttatact	aataatctta	tagaaattta	5520
ggttaaatac	agaccaagag	ccttcaaagc	cctcagtaag	ttgcaataact	taatttctgt	5580
aacagctaag	gactgcaaaa	ccccactctg	catcaactga	acgcaaatac	gccactttaa	5640
ttaagctaag	cccttacttag	accaatggga	cttaaacc	caaacactta	gttaacagct	5700
aagcaccctaa	atcaacttgc	ttcaatctac	ttctcccgcc	gccggggaaaa	aaggcgggag	5760
aagccccggc	aggtttgaag	ctgttcttc	gaatttgc	ttcaatatga	aaatcacctc	5820
ggagctggta	aaaagagggcc	taacccctgt	ctttagattt	acagtccat	gcttcactca	5880
gccattttac	ctcacccttca	ctgatgttc	ccgaccgtt	actattctt	acaaaccaca	5940
aagacattgg	aacactatac	ctattattcg	gfcgtatgac	ttggagtcc	ggcacagctc	6000
taagcctctt	tattcgagcc	gagctgggccc	agccaggcc	ccttcttagt	aacgaccaca	6060
tctacaacgt	tatcgtcaca	gccatgtcat	ttgtataat	cttcttcata	gtatatacc	6120
tcataatctgg	aggctttggc	aactgactag	ttcccttaat	aatcggtgc	cccgatatgg	6180
cgtttccccg	cataaaacaac	ataagcttct	gactttaacc	tccctcttc	ctactcctgc	6240
tcgcatctgc	tatagtggag	gcccggagcag	gaacaggtt	aacagtctac	cctcccttag	6300
cagggacta	ctcccaccc	ggagccctccg	tagacctaac	catcttctcc	ttacaccttag	6360
caggtgtctc	ctctatctta	ggggccatca	atttcatcac	aacaattatc	aatataaaaac	6420
cccctgccat	aacccaatac	caaacgcccc	tcttcgtct	atccgtcc	atcacagcag	6480
tcctacttct	cctatcttc	ccagtcctag	ctgctggcat	cactataacta	ctaacagacc	6540
gcaacctcaa	caccaccc	ttgacccccc	ccggaggagg	agacccca	ctataccaac	6600
acctattctg	attttcgg	caccctgaag	tttatattct	tatcctacca	ggcttcggaa	6660
taatctccca	tattgtact	tactactcc	aaaaaaaaga	accatttg	tacataggt	6720
ttggctgagc	tatgatatac	attggcttcc	tagggtttat	cgtgtgagca	caccatataat	6780
ttacagtagg	aatagacgt	gacacacgag	catatttcac	ctccgctacc	ataatcatcg	6840
ctatccccac	cgcggtcaaa	gtatattct	gactcgccac	actccacgga	agcaatata	6900
aatgatctgc	tgcagtgc	tgagccctag	gattcatctt	tctttcacc	gtaggtggcc	6960

Fig. 2C

tgactggcat	tgtatttagca	aactcatcac	tagacatcgt	actacacgac	acgtactacg	7020
ttgtagccca	cttccactat	gtccttatcaa	taggagctgt	atttgcacatc	ataggaggct	7080
tcattcaactg	atttccccta	ttctcaggct	acaccctaga	ccaaacacctac	gccaaaatcc	7140
atttcaactat	catattcatc	ggcgtaaatc	taactttctt	cccacaaacac	tttctcgcc	7200
tatccggaat	gccccgacgt	tactcggact	accccgtatgc	atacaccaca	tgaaacatcc	7260
tatcatctgt	aggctcattc	atttctctaa	cagcagtaat	attaataatt	ttcatgattt	7320
gagaaggcctt	cgcttcgaag	cgaaaagtcc	taatagtaga	agaaccctcc	ataaacctgg	7380
agtgactata	tggatgcccc	ccacccttacc	acacattcga	agaaccctgt	tacataaaat	7440
ctagacaaaa	aaggaaaggaa	tcgaaccccc	caaagctgtt	tcaagccaa	ccccatggcc	7500
tccatgactt	tttcaaaaag	gtatttagaa	aaccatttca	taactttgtc	aaagttaaat	7560
tataggctaa	atcctatata	tcttaatggc	acatgcagcg	caagtaggtc	tacaagacgc	7620
tacttcccct	atcatagaag	agcttateac	ctttcatgat	cacgcccctca	taatcatttt	7680
ccttatctgc	ttccttagtcc	tgtatgccct	tttccttaaca	ctcacaacaa	aactaactaa	7740
tactaacatc	tcagacgctc	aggaaataga	aaccgtctga	actatcctgc	ccgccccat	7800
ccttagtcctc	atcgccctcc	catccctacg	catccttac	ataacagacg	aggtcaacga	7860
tccctccctt	accatcaaat	caattggcca	ccaatggtac	tgaacctacg	agtacaccga	7920
ctacggcgg	ctaattttca	actcctacat	acttccccca	ttatttcctag	aaccaggcga	7980
cctgcgactc	tttgacgttg	acaatcgagt	agtactccc	attgaagccc	ccattcgtat	8040
aataattaca	tcacaagacg	tcttcactc	atgagctgtc	cccacattag	gcttaaaaac	8100
agatgcaatt	cccggacgtc	taaaccaaac	cactttcacc	gctacacgac	cggggtata	8160
ctacggtcaa	tgctctgaaa	tcttgggagc	aaaccacagt	ttcatgccc	tcgtccctaga	8220
attaattccc	ctaaaaatct	ttgaaaatagg	gcccgtattt	accctatacg	accctctcta	8280
ccccctctag	agccccactgt	aaagctaact	tagcattaac	cttttaagg	aaagattaag	8340
agaaccaaca	cctctttaca	gtgaaaatgcc	ccaactaaat	actaccgtat	ggcccaccat	8400
aattcccccc	atactcctta	cactattcct	catcacccaa	ctaaaaatat	taaacacaaa	8460
ctaccaccta	cctccctcac	caaagccccat	aaaataaaaa	aattataaca	aaccctgaga	8520
acccaaatga	acgaaaaatct	gttcgcttca	ttcattgccc	ccacaatcct	aggcctaccc	8580
gccgcgtac	tgatcatttct	atttccccct	ctattgatcc	ccacctccaa	atatctcatc	8640
aacaaccgac	taatcaccac	ccaacaatga	ctaataaacc	taacctcaaa	acaaatgata	8700
accatacaca	acactaaagg	acgaacctga	tctcttatac	tagtacccct	aatcattttt	8760
attgcaccaa	ctaaccctct	cggactcctg	cctcactcat	ttacaccaac	caccaacta	8820
tctataaacc	tagccatggc	catcccccta	tgagcggca	cagtgattat	aggcttcgc	8880
tctaagatta	aaaatgccc	agcccacttc	ttaccacaa	gcacacctac	accctttatc	8940
cccatactag	ttattatcga	aaccatcagc	ctactcattc	aaccaatagc	cctggccgt	9000
cgccttaaccg	ctaaccattac	tgcaaggccac	ctactcatgc	acctaattgg	aagcgccacc	9060
ctagcaatat	caaccattaa	ccttccctct	acacttatca	tcttcacaat	tctaattcta	9120
ctgactatcc	tagaaatcgc	tgtcgcccta	atccaaagct	acgttttcac	acttcttaga	9180
agccttacc	tgcacgacaa	cacataatga	cccaccaatc	acatgcctat	catatagtaa	9240
aaccaggccc	atgaccctca	acagggggccc	tctcagccct	cctaattgacc	tccggccctag	9300
ccatgtgatt	tcacttccac	tccataacgc	tcctcatact	aggcctacta	accaacacac	9360
taaccatata	ccaatgatgg	cgcgatgtaa	cacgagaaag	cacataccaa	ggccaccaca	9420
caccacctgt	ccaaaaaggc	cttcgatacg	ggataatcct	atttattacc	tcagaagttt	9480
ttttcttcgc	aggattttc	tgagcctttt	accactccag	cctagccct	accccccata	9540
taggaggc	ctggccccc	acaggcatca	ccccgtaaa	tcccctagaa	gtcccactcc	9600
taaacacatc	cgttattactc	gcatcaggag	tatcaatcac	ctgagctcac	catagtctaa	9660
tagaaaacaa	ccgaaaccaa	ataattcaag	cactgctt	tacaatttt	ctgggtctct	9720
attttaccct	cctacaagcc	tcaagact	tcgagctcc	tttcaccatt	tccgacggca	9780
tctacggctc	aacattttt	gtagccacag	gcttcacgg	acttcacgtc	attattggct	9840
caactttct	cactatctgc	ttcatccgcc	aactaatatt	tcactttaca	tccaaacatc	9900
actttggctt	cgaaggcc	gcgtatact	ggcatttgt	agatgtgtt	tgactatttc	9960
tgtatgtctc	catctattga	tgagggtctt	actcttttag	tataaaatagt	accgttaact	10020
tccaattaac	tagtttgac	aacattcaaa	aaagagtaat	aaacttcgcc	ttaattttaa	10080
taatcaacac	cctcctagcc	ttactactaa	taatttattac	attttgacta	ccacaactca	10140
acggctacat	agaaaaatcc	accccttacg	agtgcggctt	cgaccctata	tccccccgccc	10200
gcgtcccttt	ctccataaaaa	ttttctttag	tagcttattac	cttcttatt	tttgatctag	10260
aaattggcc	ccttttaccc	ctaccatgag	ccctacaaac	aactaacctg	ccactaatag	10320
ttatgtcatc	cctcttattta	atcatcatcc	tagcccttaag	tctggccat	gagtgactac	10380
aaaaaggatt	agactgaacc	gaattggat	atagttaaa	caaaccgaaat	gatttcgact	10440

*Fig. 2D*

cattaaatta	tgataatcat	atttacccaaa	tgccccctcat	ttacataaaat	attatactag	10500
catttaccat	ctcaacttcta	ggaataacttag	tatatcgctc	acacctcata	tcctccctac	10560
tatgcctaga	aggaaaaataa	ctatcgctgt	tcattatagc	tactctcata	accctcaaca	10620
cccactccct	cttagccaaat	attgtgccta	ttgccataact	agtctttgcc	gcctgcaag	10680
cagcggtggg	cctagcccta	ctagtctcaa	tctccaacac	atatggccta	gactacgtac	10740
ataaacctaaa	cctactccaa	tgctaaaaact	aatcgcccc	acaattatata	tactaccat	10800
gacatgactt	tccaaaaaaac	acataatttg	aatcaacaca	accacccaca	gcctaattat	10860
tagcatcatc	cctctactat	tttttaacca	aatcaacaac	aacctattta	gctgtcccc	10920
aacctttcc	tccgacccccc	taacaacccc	cctcctaata	ctaactacct	gactcctacc	10980
cctcacaatc	atggcaagcc	aacgccactt	atccagtgaa	ccactatcac	aaaaaaaact	11040
ctacctctct	atactaatact	ccctacaaat	ctccttaatt	ataacattca	cagccacaga	11100
actaatcata	tttttatatct	tcttcgaaac	cacacttac	cccaccttgg	ctatcatcac	11160
ccgatgaggc	aaccagccag	aacgctgaa	cgcaggaca	tacttcctat	tctacaccct	11220
agttaggctcc	cttcccctac	tcatcgact	aatttacact	cacaacaccc	taggctcact	11280
aaacattcta	ctactcactc	tcactgcccc	agaactatca	aactcctgag	ccaataactt	11340
aatatgacta	gcttacacaa	tagcttttat	agtaaagata	cctcttttag	gactccactt	11400
atgactccct	aaagccccatg	tcgaagcccc	catcgctggg	tcaatagtac	ttgcccaggt	11460
actcttaaaa	ctaggcggt	atggtataat	acgcctcaca	ctcattctca	acccctgac	11520
aaaacacata	gcctacccct	tccttgtact	atccctatga	ggcataatta	taacaagctc	11580
catctgccta	cgacaaacacag	acctaaaatc	gctcattgca	tactcttcaa	tcagccacat	11640
agccctcgta	gtaacagcca	ttctcatcca	aacccctga	agttcaccg	gcbcagtcat	11700
tctcataatc	gccccacgggc	ttacatccctc	attactattc	tgccctagcaa	actcaaacta	11760
cgaacgcact	cacagtgcga	tcataatct	ctctcaagga	cttc当地actc	tactcccact	11820
aatagttt	tgatgacttc	tagcaaggct	cgctaaactc	gccttacccc	ccactattaa	11880
cctactggga	gaactctctg	tgcttagtaac	cacggttctcc	tgatcaaata	tcactctcct	11940
acttacagga	ctcaacatac	tagtcacagc	cctataactcc	ctctacatat	ttaccacaac	12000
acaatggggc	tcactcaccc	accacattaa	caacataaaa	ccctcattca	cacgagaaaa	12060
caccctcatg	ttcatacaccc	tatccccat	tctcctcta	tccctcaacc	ccgacatcat	12120
taccgggtt	tcctcttgt	aatatagtt	aacccaaaaca	tcaagattgt	aatctgacaa	12180
cagaggctta	cgacccctta	tttaccgaga	aagctcacaa	gaactgctaa	ctcatgcccc	12240
catgtctaac	aacatggctt	tctcaacttt	taaaggataa	caagtatcca	ttggtcttag	12300
gccccaaaaaa	ttttggtgca	actccaaata	aaagtaataa	ccatgcacac	tactataacc	12360
accctaaaccc	tgacitccct	aattcccccc	atccttacca	ccctcgtaaa	ccctaacaaaa	12420
aaaaactcat	accccccatta	tgtaaaatcc	attgtcgat	ccacctttat	tatcagtctc	12480
ttccccacaa	caatattcat	gtgccctagac	caagaagtt	ttatctcgaa	ctgacactga	12540
gccacaaccc	aaacaaaccca	gcttcccta	agttcaaaac	tagactactt	ctccataata	12600
ttcattccctg	tagcattgtt	cgttacatgg	tccatcatag	aattctcact	gtgatata	12660
aactcagacc	caaacattaa	tcagttcttc	aaatatctac	tcatcttcct	aattaccata	12720
ctaattctag	ttaccgctaa	caacctattc	caactgttca	tcggctgaga	ggcgtagga	12780
attatatctt	tcttgctcat	cagttgatga	tacgcccag	cagatgcca	cacagcagcc	12840
attcaagcaa	tcctataacaa	ccgtatcgcc	gatatcggtt	tcatctcgc	cttagcatga	12900
tttattctac	actccaactc	atgagaccca	caacaaataq	cccttctaaa	cgctaattcca	12960
agcctcaccc	cactactagg	cctccctcta	gcagcagcag	gcaaattcagc	ccaatttaggt	13020
ctccacccct	gactcccttc	agccatagaa	ggccccaccc	cagtctcagc	cctactccac	13080
tcaagacta	tagttgtac	aggaatcttc	ttactcatcc	gcttccaccc	cctagcagaa	13140
aatagcccac	taatccaaac	tctaacaacta	tgcttaggc	ctatcaccac	tctttcgca	13200
gcagtcg	ccttacaca	aaatgacatc	aaaaaaatcg	tagccttctc	cactcaagt	13260
caactaggac	tcataatagt	tacaatcgcc	atcaaccaac	cacacctagc	attcctgcac	13320
atctgtaccc	acgccttctt	caaagccata	ctatttatgt	gctccgggtc	catcatccac	13380
aaccttaaca	atgaacaaga	tattcgaaaa	ataggaggac	tactcaaaac	catacctctc	13440
acttcaaccc	ccctcaccat	tggcagccta	gcatttagcag	gaaatactt	cctcacaggt	13500
ttctactcca	aagaccacat	catcgaaacc	gcaaacatat	catacaca	cgccctgagcc	13560
ctatctatta	ctctcatcg	tacccctctg	acaaggcct	atagcactcg	aataattctt	13620
ctcacccctaa	caggtcaacc	tcgcttcccc	acccttacta	acattaacga	aaataacccc	13680
accctactaa	accccataa	acgcctggca	gcccggaaagcc	tattcgcagg	atttctcatt	13740
actaacaaca	tttccccccgc	atcccccttc	caaacaacaa	tcccctctca	cctaaaactc	13800
acagccctcg	ctgtcacttt	cctaggactt	ctaacagccc	tagacctcaa	ctacctaacc	13860
aacaaactta	aaataaaaatc	cccactatgc	acattttatt	tctccaaacat	actcgattc	13920

6/42

taccctagca	tcacacacccg	cacaatcccc	tatctaggcc	ttcttacgag	ccaaaacctg	13980
cccctactcc	tcctagacct	aacctgacta	gaaaagctat	tacctaaaac	aatttcacag	14040
caccaaatct	ccacctccat	catcacctca	acccaaaaag	gcataattaa	actttacttc	14100
ctctctttct	tcttcccact	catcctaacc	ctactcctaa	tcacataacc	tattccccg	14160
agcaatctca	attacaatat	atacaccaac	aaacaatgtt	caaccagtaa	ctactaactaa	14220
tcaacggcca	taatcataca	aagccccgc	accaatagga	tcctcccgaa	tcaaccctga	14280
ccccctctct	tcataaaatta	ttcagcttcc	tacactatta	aagtttacca	caaccaccac	14340
cccatcatac	tctttcacc	acagcaccaa	tcctacctcc	atcgctaacc	ccactaaaac	14400
actcaccagg	acctcaaccc	ctgacccccc	tgcctcagga	tactcctcaa	tagccatcgc	14460
tgttagtat	ccaaagacaa	ccatcattcc	ccctaaataaa	ataaaaaaaa	ctattaaacc	14520
catataac	cccccaaaat	tcagaataat	aacacacccg	accacacccg	taacaatcaa	14580
tactaaaccc	ccataaaatag	gagaaggctt	agaagaaaaac	cccacaaaacc	ccattactaa	14640
acccacactc	aacagaaaaca	aagcatacat	cattatttc	gcacggacta	caaccacgac	14700
caatgatatg	aaaaaccatc	gttgtatttc	aactacaaga	acaccaatga	ccccaaatacg	14760
caaaattaac	cccctaataa	aattaattaa	ccactcattc	atcgacactcc	ccaccccatc	14820
caacatctcc	gcatgatgaa	acttcggctc	actccttggc	gcctgcctga	tcctccaaat	14880
caccacagga	ctattccctag	ccatgcacta	ctcaccagac	gcctcaacccg	ccttttcatc	14940
aatcgcccac	atcactcgag	acgtaaattaa	tggctgaatc	atccgctacc	ttcacgccaa	15000
tggcgccctca	atattcttta	tctgccttct	cctacacatc	gggcgaggcc	tatattacgg	15060
atcatttctc	tactcagaaa	cctgaaacat	cggcattatc	ctcctgcttg	caactatagc	15120
aacagccttc	ataggctatg	tcctcccg	aggccaaata	tcattctgag	gggcccacagt	15180
aattacaaac	ttactatccg	ccatcccata	cattgggaca	gacctagttc	aatgaatctg	15240
aggaggctac	tca	gtcccccac	cacacgattc	tttaccttcc	acttcatctt	15300
gcccttcatt	attgcagccc	tagcaacact	ccacctccta	ttcttgac	aaacgggatc	15360
aaacaacccc	cttaggaatca	cctcccatc	cgataaaatc	acttccacc	cttactacac	15420
aatcaaagac	gccctcggt	tacttcttct	ccttctctcc	ttaatgacat	taacactatt	15480
ctcaccagac	ctcctaggcg	acccagacaa	ttatacccta	gccaacccct	taaacacccc	15540
tccccacatc	aagcccgaat	gatatttcct	attcgctac	acaattctcc	gatccgtccc	15600
taacaaacta	ggaggcg	ttgcctatt	actatccatc	ctcactcctag	caataatccc	15660
catcctccat	atatccaaac	aacaaagcat	aatatttgc	ccactaagcc	aatcaactta	15720
ttgactccta	gccgcagacc	tcctcattct	aacctgaatc	ggaggacaac	cagtaagcta	15780
ccctttacc	atcattggac	aagttagcatc	cgtactatac	ttcacaacaa	tcctaatcct	15840
aataccaact	atctccctaa	ttgaaaacaa	aataactcaa	tgggcctgtc	ctttagtat	15900
aaactaatac	accagtcttq	taaaccggag	atgaaaacct	tttccaagg	acaatcaga	15960
aaaaaagtct	ttaactccac	cattagcacc	caaagctaag	attctaattt	aaactattct	16020
ctgttcttc	atggggaa	agattttgggt	accacccaag	tattgactca	cccatcaaca	16080
accgctatgt	atttcgta	ttactgccc	ccaccatgaa	tattgtacgg	taccataaat	16140
acttgaccac	ctgtgttaca	taaaaaaccca	atccacatca	aaacccccc	cccatgctta	16200
caagcaagta	cagcaatcaa	ccctcaacta	tcacacatca	actgcaactc	caaagccacc	16260
cctcacccac	taggatacc	acaaacctac	ccacccttaa	cagtagatag	tacataaagc	16320
catttaccgt	acatagcaca	ttacagtcaa	atcccttctc	gtccccatgg	atgacccccc	16380
tcagataggg	gtcccttgac	caccatc	cgtgaaatca	atatcccgca	caagagtgt	16440
actctccctcg	ctccggggcc	ataacacttg	gggtagcta	aagtgaactg	tatccgacat	16500
ctggttccta	ttcagggtc	ataaaagccta	aatagccac	acgttccct	taaataagac	16560
	atcacatg.					16569

*Fig. 2E*

*Fig. 3A*

ataggcattc	caactcatcc	gcttgctgac	gacctccacg	tgatttcaac	aatgatttca	60
aatatttcac	tttttaagtgc	agtgtgactt	aagtatgaaa	ttgcctctcc	ctaaagctcc	120
cctaaggcct	aaacagtgcgt	cattaccata	gctgtgacag	ggagactgtt	gaatttataa	180
tctattggcc	attcacagca	tagcgtataa	accctagctc	atgatttctt	tgcaatagaa	240
gtgtactttt	tcatcacatt	cccttcacaa	cttactcacc	agacagactt	tgagctctcc	300
tcctggctta	gcctggatcg	tttggaaatgg	tcatccatcc	tttggccccca	atacctaaac	360
taaggcttat	gaacaataaag	atgattttac	ttcagtggg	cttttttgtt	taatatattha	420
gatttgacct	tcagcaaggt	caaagggagt	ccgaactagt	ctcaggcttc	aacatcgaat	480
acgcccgcagg	ncccttcgcc	ctattcttca	tagccgataa	cacaaacattt	attataataa	540
acaccctcac	caactacaatc	ttccttaggaa	caacatataa	cgcactctcc	cctgaactct	600
acacaacata	tttgtcacc	aagaccctac	ttctgacctc	cctgttctta	tgaattcgaa	660
cagcataccc	ccgattccgc	tacgaccaac	tcatacacct	cctatgaaaaa	aacttcctac	720
cactcaccct	agcattactt	atatgatatg	tctccatacc	cattacaatc	tccagcattc	780
cccctcaaac	ctaagaataa	tgtctgataa	aagagttact	ttgatagagt	aaataatagg	840
agtttaaattc	cccttatttc	taggactatg	agaatcgaac	ccatccctga	gaatccaaaaa	900
ttctcggtgc	caactatcac	acccatctt	aaagtaaggt	cagctaaata	agctatcgaa	960
cccatacccc	gaaaatgtt	gttatatcct	tccgtaacta	attaatcccc	tggcccaacc	1020
cgtcatttac	tctaccatct	ttgcaggcac	actcatcaca	gcgctaagct	cgcactgatt	1080
ttttacctga	gtaggcctag	aaataaacat	gctagtttt	attccagttc	taaccaaaaaa	1140
aataaaccct	cgttccacag	aagntgccat	caagtatttc	ctcacgcaag	caaccgcata	1200
cataatcctt	ctaatacgta	tcctnttcaa	caatatactc	tccggacaat	gaaccatwac	1260
caataccacc	aatcaatact	catcattaat	aatcataatg	gctatagcaa	taaaactagg	1320
aatagcccccc	tttcacttct	gagtcccaga	ggttacccaa	ggcacccctc	tgacatccgg	1380
cctgctccctt	ctcacatgac	aaaaactagc	ccccatctca	atcatataacc	aaatttctcc	1440
ctcattaaac	gtaaggcctt	tcctcactct	ttcaatctta	tccatcatgg	caggcagtgt	1500
aggtggatta	aaccaaaaacc	aactacgcaa	aatcttagca	tactcctcaa	ttacccacat	1560
aggatgaata	acagcaggtc	tacggtacaa	ccctaacata	accattctta	attnaactat	1620
ttatattatc	ctaactacta	ccgcattccct	actactcaac	ttaaactcca	gcaccacaac	1680
cctactacta	tctcgacact	gaaacaagct	aacatgacta	acacccttaa	ttccatccac	1740
cctccctctcc	ctaggaggcc	tgcccccgct	aaccgcttt	ttgccccaaat	gggcccattat	1800
cgaagaattc	acaaaaaaaca	atagcctcat	catccccacc	atcatagcca	tcatcaccct	1860
ccttaacctc	tacttctacc	tgcgcctaat	ctactccacc	tcaatcacac	tactccctat	1920
atctaacaac	gtaaaaataa	aatgacagtt	tgaacacaca	aaaccacccc	cattcctccc	1980
cacactcatc	gcccttacca	cactgctccf	acctatctcc	ccttttatgc	taataatctt	2040
atagaaaattt	aggftaaata	cagaccaaga	gccttcaaag	ccctcagtaa	tttgcaatac	2100
ttaatttctg	naacagctaa	ggactgcaaa	accccaactct	gcatcaactg	nangcaaata	2160
agccacttta	attaagctaa	gcccttacta	gaccaatggg	acttaaacc	acaacacatt	2220
agttAACAGC	taagcaccc	aatcaactgg	cttcaatcta	cttctcccg	cgccgggaaa	2280
aaaggcggga	gaagccccgg	caggtttgaa	gctgcttctt	cgaatttgca	attcaatatg	2340
aaaatcacct	canagctgt	aaaaagaggc	ttaaccctg	tcttttagatt	tacagtccaa	2400
tgcttcactc	agccatttt	cctcaccccc	actgatgttc	gccgaccgtt	gactattctc	2460
tacaaaccac	aaagacattt	gaacactata	ccttattttc	ggcgcatgag	ctggagtctt	2520
aggcacagct	ctaagccctcc	ttattcgac	cgaactggc	cagccaggca	accttcttagg	2580
taacgaccac	atctacaacg	ttatcgatc	agcccattgc	tttgtaataa	tcttcttcat	2640
agtaataaccc	atcataatcg	gaggytttg	caastgacta	gttcccctaa	taatcggtgc	2700
ccccgatatg	gsgtttcccc	gcataaaca	cataagctt	tgactcttac	ccccctctct	2760
cntactcctg	ttnngcatctg	ctatagtgg	ggccggcgca	ggaacagggt	gaacagtna	2820
ccctcccttg	gcagggact	actcccaccc	tggagctcc	gtagacctaa	ccatcttctc	2880
cttacaccta	gcaggtatct	cctctatctt	agggaccatc	aatttcatca	caacaattat	2940
taatataaaa	ccccctgcca	taacccaata	ccaaacgccc	ctttctgtt	gatccgtcct	3000
aatcacagca	gtcttacttc	tcctatctct	cccagtctta	gccgctggca	tcactatact	3060
actaacagac	cgtaacctca	acaccacctt	cttcgaccca	gccggagggag	gagacccat	3120
tctataccaa	cacctattct	gattttcgg	tcacccctgaa	gtttatattc	tcatcttacc	3180
aggcttcgga	ataatctccc	atattgtaac	ttactactcc	ggaaaaaaaaaag	aaccatttgg	3240
atacataggt	atggctctgag	ctatgatatac	aattggcttc	ctagggttta	tcgtgtgagc	3300
acaccatata	tttacagtag	gaatagacgt	agacacacga	gcatatttca	cctccgtctac	3360
cataatcatc	gctatcccc	ccggcgtaaa	agtatttagc	tgactcgcca	cactccacgg	3420
aagcaatatg	aatgatctg	ctgcagtgct	ctgagcccta	ggattttattt	ttcttttcac	3480

8/42

cgtagggtggc	ctgactggca	ttgtattagc	aaactcatca	ctagacatcg	tactacacga	3540
cacgtactac	gttgtagccc	acttccacta	tgtcctatca	ataggagctg	tatttgc当地	3600
cataggaggc	ttcattcact	gattttccct	attctcaggc	tacaccctag	accaaaccct	3660
cgc当地aaatc	catttcgcta	tcatattcat	cgcgtaaat	ctaactttct	tcccacaaca	3720
ctttctcggc	ctatccggaa	tgccccgacg	ttactcggac	tatcccgtg	catacaccac	3780
atgaatatac	ctatcatctg	taggctcatt	catttctct	acagcagtaa	tattaataat	3840
tttcataatt	tgagaaggct	tcgcttcgaa	gcgaaaagtc	ctaatacgtag	aagaaccctc	3900
cataaaacctg	gagtgactat	atggatgccc	cccaccctac	cacacattcg	aagaaccctg	3960
atacataaaaa	tctagacaaaa	aaaggaagga	atcgaaacccc	ccaaagctgg	tttcaagcca	4020
accccatggc	ctccatgact	ttttcaaaaa	gatattagaa	aaaccatttc	ataactttgt	4080
caaagtaaa	ttataggcta	aatccatat	atcttaatgg	cacatgcagc	gcaagtaggt	4140
ctacaagacg	ctacttcccc	tatcatagaa	gagtttatca	tctttcatga	tcacgcccctc	4200
ataatcattt	tccttatctg	tttcctagtc	ctgtacgccc	ttttcctaac	actcacaaca	4260
aaactaacta	atactaacat	ctcagacgct	caggaaatag	aaaccgtctg	aactatcctg	4320
cccgccatca	tccttagtct	tatgcctc	ccatnccctac	gcatccttta	cataacagac	4380
gaggtcaacg	atcccctcctt	taccatcaaa	tcaattggcc	atcaatggta	ctgaacctac	4440
gaatacacccg	actacggcgg	actaatcttc	aactcctaca	tacttcccc	attattccta	4500
gaaccaggcg	acctgcgact	ccttgcgtt	gacaatcgag	tagtactccc	ggttgaagcc	4560
cccattcgta	taataattac	atcacaagac	gtcttacact	catgagctgt	ccccacacatta	4620
ggcttaaaaa	cagatgcaat	tcccggacgt	ctaaacccaaa	ccactttcac	tgctacacga	4680
ccaggggtat	actacggcca	atgctctgaa	atctgtggag	caaaccagg	ttatgccat	4740
cgtcctagaa	ttaattcccc	taaaaatctt	tgaaataggg	cccgatttta	ccctatagca	4800
ccccctctac	ccccctctaga	gcccactgt	aagctaattt	agcattaacc	ttttaagtta	4860
aagattaaga	gaaccaacac	ctctttacag	tgaaatgccc	caactaaata	ctaccgtatg	4920
acccaccata	attaccccca	tactccttac	actattctc	atcaccacac	taaaaatatt	4980
aaatacaaaat	taccacctac	ctccctcacc	aaagccata	aaaataaaaa	actataacaa	5040
accctgagaa	ccaaaatgaa	cgaaaatctg	ttcacttcat	tcattgcccc	cacaatccta	5100
ggcctaccgg	ccgcgactact	gatcattcta	ttyccccctc	tattgatccc	caccccaaaa	5160
tatctcatca	acaaccgact	aattaccacc	caacaatgac	taatccaact	aacctaaaa	5220
caaatgatag	ccatacacaa	cactaaggga	cgaacctgat	ctcttataact	agtatcctt	5280
atcatttta	ttgccacaac	taaccccttc	ggactctgtc	ctcactcatt	tacaccaacc	5340
acccaactat	ctataaaact	agccatggcc	atccccctat	gagcgggccc	agtgattata	5400
ggcttcgct	ctaaagattaa	aaatgcccct	gcccacttct	taccacaagg	cacacctaca	5460
ccccttatcc	ctataactgt	tattatcgaa	accatcagcc	tactcattca	accaatagcc	5520
ctggccgtan	gnctaaccgc	taacattact	gcagggccacc	tactcatgna	cctaattgga	5580
agcgccacac	tagcaatatc	aantattaac	cttccctcta	cacttacat	tttcacaatt	5640
ctaattctac	tgactatcc	agaaatcgct	gtcgccctaa	tccaagccta	cgttttaca	5700
cttttagtaa	gcctctacct	gcacgacaac	acataatgac	ccaccaatca	catgcctatc	5760
atatagtaaa	acccagccca	tggcccctaa	cagggccct	ccttataacta	ggcctactaa	5820
ccggcctagc	catgtgattt	cacttccact	ccacaaccct	ccttataacta	acataccaag	5880
ccaacacact	aaccatatac	caatgatggc	gcgatgtaac	acgagaaagc	acataccaag	5940
gccaccacac	accacctgtc	cagaaggccc	ttcgatcgg	gataatccta	tttattacct	6000
ccccccaact	agggggacac	tggcccccac	cagggcatcac	cccgctaaat	ccccctagaag	6120
tcccactcct	aaacacatcc	gtattactcg	catcaggggt	atcaatcacc	tgagctcacc	6180
atagtctaatt	agaaaacaac	cgaaaacaaa	taattcaagc	actgcttatt	acaattttac	6240
tgggtctcta	ttttaccctc	ctacaaggct	cagngtactt	cgaggttaaa	atattagata	6300
tttcccctga	tacagggttc	aatcttttc	tttttaaagc	aatatttctc	aaagtacttt	6360
tcacagaact	taagtttcat	taagcacttc	actaaaaagna	aaagtctgt	atctaataaa	6420
tttggaaaat	attgagaatt	agagccccct	cttagatatg	tactgtatg	actcagcttg	6480
ttacagatgg	aagtaaacat	tgtaatattc	acccagctt	tgagtggatg	tcttattaaca	6540
tcacccaaat	gagtattcca	tggaaatgcac	tttgcaaaaa	ccttattttc	aagaaaaatt	6600
ctggagcatg	gaaagctatt	aatggataaa	cccattcaca	aaatcacacc	aaatatctaa	6660
aatcatgttt	aaaatctcct	agaaatgggt	t			6691

*Fig. 3B*

*Fig. 4A*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito-chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
COI	5.8 kb pg	6023	G	A
		6221	T	C
		6242	C	T
		6266	A	CG
		6299	A	G
		6366	G	A
		6383	G	AT
		6410	C	T
		6452	C	TCT
		6483	C	TA
		6512	T	CT
		6542	C	ACT
		6569	C	TTGTT
		6641	T	T
		6935	C	A
		6938	C	T
		7146	A	A
		7232	C	T
		7256	C	T
		7316	G	A
COI	11	6023	G	A
COI	12	6221	T	CTC
		6242	C	CG
		6266	A	G
		6299	A	A
		6366	G	AT
		6383	G	T
		6410	C	
COI	12	6160	T	A
		6182	G	C
		6185	T	C
		5216	T	C
		6221	T	CTT
		6224	C	T
		6236	C	TC
		6242	C	T
		6251	T	C
		6260	G	A
		6266	A	CCG
		6269	A	GT
		6281	A	T
		6326	C	TG
		6335	C	T
		6353	A	CAC
		6356	A	T
		6365	T	C
		6366	G	A
		6378	T	C

*Fig. 4B*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito- chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
COI	12	6383	G	A
		6389	C	T
		6392	T	C
		6398	C	T
		6407	T	C
		6410	C	T
		6260	G	A
		6266	A	CC
		6269	A	CG
		6281	A	TTGTC
		6326	CC	TTGTC
		6335	CA	AC
COI	12	6353	ACT	ACT
		6356	CT	AC
		6365	GT	A
		6366	GT	AC
		6378	GT	AC
		6383	GT	A
		6392	TC	CT
		6398	C	T
		6182	G	A
		6216	TT	CC
		6221	GA	CCG
		6260	AA	TTGTC
COI	12	6266	AA	AC
		6269	AA	CG
		6281	AC	TTGTC
		6326	CC	AC
		6335	CA	GT
		6353	ACT	TC
		6356	CT	AC
		6365	GT	AC
		6366	GT	AC
		6378	GT	AC
		6383	GT	A
		6392	TC	CT
COI	12	6398	C	T
		6160	T	A
		6182	G	AC
		6185	TT	CC
		5216	TT	CTT
		6221	TC	CTT
		6224	CC	TT
		6236	CT	TC
		6242	CT	CC
		6251	TA	CC
		6266	A	CC
		6269	A	CC

*Fig. 4C*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito- chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		6326	C	T
		6335	C	T
		6353	A	G
		6356	C	T
		6365	T	C
		6366	G	A
		6378	T	C
		6383	G	A
		6392	T	C
COI	12	6160	T	A
		6182	G	A
		6185	T	C
		5216	T	C
		6221	TC	TT
		6224	CT	TC
		6242	TG	AC
		6251	AA	CC
		6260	AC	GT
		6266	AT	TT
		6269	AG	TC
		6281	AT	GA
		6326	AC	CT
		6335	CC	TC
		6353	AA	GT
		6356	CT	TC
		6365	TT	CA
		6366	GT	AC
		6378	TT	CA
		6383	GC	AT
		6389	TC	TC
		6392	CT	TC
		6398	CT	TC
		6407	TC	CT
		6410	C	T
COI	12	6251	T	C
		6260	G	A
		6266	A	C
		6269	A	CG
		6281	AC	TT
		6326	C	GT
		6335	CC	TC
		6353	AA	GT
		6356	CT	TC
		6365	TT	CA
		6366	GT	AC
		6378	TT	CA
		6383	GC	AT
		6392	T	AC

*Fig. 4D*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito-chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		6398	C	T
COI	12	6185	T	C
		6251	T	C
		6260	G	A
		6266	A	C
		6269	A	C
		6281	A	G
		6326	C	T
		6335	C	T
		6353	A	G
		6356	C	T
		6365	T	C
		6366	G	A
		6378	T	C
		6383	G	A
		6392	T	C
		6398	C	T
COI	12	6185	T	C
		6236	C	T
		6251	T	C
		6260	G	A
		6266	A	C
		6269	A	C
		6281	A	G
		6326	C	T
		6335	C	T
		6353	A	G
		6356	C	T
		6365	T	C
		6366	G	A
		6378	T	C
		6383	G	A
		6392	T	C
		6398	C	T
COI	12	6236	C	T
		6251	T	C
		6260	G	A
		6266	A	C
		6269	A	C
		6281	A	G
		6326	C	T
		6335	C	T
		6353	A	G
		6356	C	T
		6365	T	C
		6366	G	A
		6378	T	C
		6383	G	A

*Fig. 4E*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito- chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		6392	T	C
		6398	C	T
COI	12	6236	C	T
		6251	T	C
		6260	G	A
		6266	A	C
		6269	A	CCG
		6281	A	CGT
		6326	C	T
		6335	C	TTG
		6353	A	G
		6356	C	TCA
		6365	T	C
		6366	G	A
		6378	T	C
		6383	G	A
		6392	T	C
		6398	C	T
COI	12	6185	T	C
		6236	C	TC
		6251	T	CA
		6260	G	ACC
		6266	A	CGT
		6269	A	G
		6281	A	TCA
		6326	C	ACT
		6353	A	G
		6356	C	TCA
		6365	T	A
		6366	G	C
		6383	G	A
		6392	T	C
		6398	C	T
COI	12	6185	T	C
		6236	C	TC
		6251	T	CA
		6260	G	ACC
		6266	A	CGT
		6269	A	G
		6281	A	TCA
		6326	C	ACT
		6353	A	G
		6356	C	TCA
		6365	T	A
		6366	G	C
		6383	G	A
COI	12	6185	T	C
		6251	T	C
		6260	G	A

*Fig. 4F*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito-chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		6266	A	C
		6269	A	C
		6281	A	G
		6326	C	T
		6335	C	T
		6353	A	G
		6356	C	T
		6365	T	C
		6366	G	A
		6383	G	A
		6392	T	C
		6398	C	T
COI	12	6185	T	C
		6251	T	C
		6260	G	A
		6266	A	CCG
		6269	A	GTT
		6281	A	TGT
		6326	C	TGTC
		6335	C	ACT
		6353	A	T
		6356	C	T
		6365	T	C
		6366	G	A
		6392	T	ACT
		6398	C	T
COI	13	6452	C	T
		6483	C	TCT
		6512	T	C
		6542	C	TACT
		6569	C	T
		6641	T	ACTT
		6935	C	T
		6938	C	T
COI	15	7146	A	G
COI	15	6950	C	A
		6962	G	A
		6965	T	GCT
		6975	T	TAC
		6990	C	ACC
		7013	G	CGT
		7040	T	T
		7064	A	CGT
		7076	T	G
		7145	A	A
		7146	C	T
		7160	A	G
		7175	T	A

*Fig. 4G*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito- chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
COI	15	6950	C	A
		6962	G	A
		6965	T	G
		6975	T	C
		7013	G	A
		7040	T	C
		7064	T	C
		7076	A	G
		7145	C	T
		7146	A	G
		7160	C	A
		7175	T	C
COI	15	6932	A	G
		6935	C	T
		6938	C	C
		6944	T	T
		6950	C	A
		6962	G	A
		6965	T	G
		6975	T	C
		6990	C	T
		7013	G	A
		7022	T	T
		7028	C	C
		7037	T	T
		7040	C	C
		7064	T	T
		7072	C	C
		7079	T	T
		7100	C	G
		7112	A	T
		7145	C	T
		7146	A	G
		7160	C	A
		7169	T	C
		7175	T	C
COI	15	6932	A	G
		6935	C	T
		6938	C	C
		6944	T	T
		6950	C	A
		6962	G	A
		6965	T	G
		6975	T	C
		6990	C	T
		7013	G	A
		7022	T	T
		7028	C	T

*Fig. 4H*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito-chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		7037	C	T
		7064	T	C
		7072	T	C
		7079	C	T
		7100	A	G
		7112	C	T
		7145	C	T
		7146	A	G
		7160	C	A
		7169	T	C
		7175	T	C
COI	15	6929	A	G
		6938	C	T
		6944	T	C
		6950	C	A
		6956	T	C
		6962	G	A
		7013	G	A
		7022	T	C
		7040	T	C
		7064	T	C
		7076	A	G
		7133	C	T
		7145	C	T
		7146	A	G
		7169	T	C
COI	15	6929	A	G
		6938	C	T
		6944	T	C
		6950	C	A
		6956	T	C
		7013	G	A
		7022	T	C
		7040	T	C
		7064	T	C
		7076	A	G
		7133	C	T
		7145	C	T
		7146	A	G
		7169	T	C
COI	15	6929	A	G
		6938	C	T
		6944	T	C
		6950	C	A
		6956	T	C
		6962	G	A
		7013	G	A
		7022	T	C

*Fig. 4I*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito- chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		7040	T	C
		7064	T	C
		7076	A	G
		7133	C	T
		7146	A	G
		7169	T	C
COI	15	6950	C	A
		6962	GT	AA
		6965	TT	AG
		6975	TC	CT
		6990	CG	TAC
		7013	TT	CCG
		7040	TA	GT
		7064	CA	TG
		7076	CT	AC
		7145	ACT	G
		7146	AT	T
		7160	CT	A
		7175	T	C
COI	15	6932	A	G
		6938	CC	TA
		6950	CG	AGC
		6962	TT	TACT
		6965	TC	CC
		6975	TC	CT
		6990	GT	G
		7013	TC	T
		7022	CC	G
		7028	CT	T
		7037	TT	C
		7040	TC	C
		7064	TT	CT
		7072	TC	T
		7079	CA	G
		7100	AC	A
		7112	CA	T
		7146	AC	G
		7160	CT	A
		7169	TT	C
		7175	T	C
COI	16	7232	C	T
		7256	CC	T
		7316	G	A
COII	5.8 kd pg	7650	CT	T
		7705	CT	C
		7810	CCC	T
		7868	CC	T
		7891	C	T

*Fig. 4J*

Human exmtDNA Nucleotide Substitutions and Deletions Relative to  
Corresponding Human mtDNA Sequence [SEQ ID NO:2]

Mito-chondrial Gene Region	Fragment	Human mtDNA Nucleotide Position No.	Human mtDNA SEQ ID NO:2	Human exmtDNA Substitution (or Deletion)
		7912	G	A
		8021	A	G
		8065	G	A
		8140	C	T
		8152	G	A
		8167	T	C
		8196-8197	AC	deletion
		8203	C	T
COII	21/22/23	7650	C	T
		7705	T	C
		7810	C	T
		7868	CC	TT
		7891	CG	TT
		7912	G	A
		8021	AG	G
		8065	GC	AT
		8140	GT	A
		8152	GT	A
		8167	T	C
		8196-8197	AC	deletion
		8203	C	T
COII	21/22/23	7650	C	T
		7705	T	C
		7810	C	T
		7868	CC	TT
		7891	CG	TT
		7912	G	A
		8021	A	G
COII	21	7650	C	T
		7705	TC	C
		7810	C	T
COII	21	7650	C	T
COII	21/22/23	7650	C	T
		7705	TT	C
		7861	T	CT
		7868	CC	TT
		7891	CC	TT
		7900	CG	TA
		7912	GC	T
		7927	A	G
		8011	AA	G
		8021	AT	G
		8038	T	C
COII	21/22/23	7650	C	T
		7663	C	TC
		7705	T	C
		7810	C	T
		7868	C	T